Screening and selection: Applications for vitamin biosynthesis in Escherichia coli

Buerger, Josephine

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Screening and selection: Applications for vitamin biosynthesis in *Escherichia coli*

Ph.D. Thesis

Josi Buerger

The Novo Nordisk Center for Biosustainability
Biosyntia ApS

January 2019

Supervisors:
Prof Morten O. A. Sommer
Dr Luisa Gronenberg

Preface

The work presented in this Ph.D thesis was performed between October 2015 and September 2018 at the Novo Nordish Center for Biosustainability at the Technical University of Denmark as well as at Biosyntia ApS, Copenhagen. Further, a six week research stay was performed at FGen FmbH in Basel, Switzerland.

The work was supervised by Prof Morten Otto Alexander Sommer (academic supervisor) and Dr Luisa Gronenberg (secondary supervisor). Dr Hans Jasper Genee also supervised the project.

Funding was generously obtained from the Horizon 2020 project as part of the Marie Skłodowska-Curie ITN “metaRNA”.

The thesis committee is composed of Dr Hans-Peter Hohmann, Dr Rute Neves (Chr. Hensen), and Dr Pablo Ivan Nikel (Biosustain, DTU).

Acknowledgements

The work presented in this thesis would not have been possible without the generous guidance and support of my supervisors Dr Luisa Gronenberg, Dr Hans Jasper Genee, and Prof Morten Otto Alexander Sommer. I would also like to thank the team at Biosyntia who have made the last three years enjoyable. Their work remains an inspiration. Finally, the Fellows and PIs of my funding network, “metaRNA”, helped guide me through this process.

Table of Contents

Abstract

The chemicals industry provides the foundation of our modern lifestyle, but it has far reaching consequences for the planet. It is no longer insider knowledge that industrial biotechnology can offer more sustainable solutions to production via fermentative biosynthesis. Nevertheless, acceptance of fermentation processes is contingent on a single factor: cost-effectiveness. Therein lies the challenge of bacterial strain engineering: to generate a cell factory capable of producing the target molecule more efficiently than a highly optimised chemical process.

Rerouting and redesign of bacterial metabolism has become a complex skill, yet rational strain engineering is insufficient to garner the required concentrations of target molecule. Random and permutatory mutagenesis can be utilised to find mutations that will increase target titers but come with a new challenge: testing thousands of mutated strains for production capacity.

In this thesis, we construct systems that can offer solutions to the “problem of screening” for biotin biosynthesis in *Escherichia coli*. After a brief introduction to our target molecule biotin with particular emphasis on papers published in the last two years, we review the relevant literature on screening and selection. We then develop a fluorescent output in *E. coli* cells based on biotin availability, and adapt it for co-encapsulation in alginate beads. Further, we reroute biotin metabolism in such a way that strain can only survive if high concentrations of the target molecule are available and use this to identify high production variants. Finally, we address the problem of false positives with a targeted construct to sieve out evolutionary cheaters.

Dansk abstrakt

Selvom kemiindustrien er grundlaget for vores moderne livsstil, har den negative konsekvenser for verden. Det *er* muligt at erstatte kemisk syntese med fermenteringsprocesser på en mere bæredygtig måde. Så længe overskud er den vigtigste betingelse for produktionen, er det dog svært at forandre industrien. Heri ligger udfordringen for industriel bioteknologi: at skabe cellefabrikker der kan producere kemiske produkter med højere effektivitet end allerede optimerede kemiske processer.

At omdirigere bakteriers metabolisme er blevet et modent forskningsfelt, men kun at gøre dette på en rationel fremgangsmåde er ikke godt nok til at nå de høje produkt koncentrationer der er brug for. Mutagenese kan også udføres tilfældigt eller via permutationer, men dette går hånd i hånd med et andet problem: tusindvis af nye stammer skal have deres produktionskapacitet testet.

Denne tese handler om at skabe systemer det kan løse dette screeningproblem, især med hensyn til biotin biosyntese i *Escherichia coli*. Der bliver indledt med en opsummeringer af den publiceret litteratur. Derefter beskriver vi et fluorescerende konstrukt i *E. coli* som afhænger at biotin koncentration og som gør det muligt at bruge denne biosensor i alginatkugler til high-throughput screening. Derudover omdirigerer vi biotin metabolismen, så stammer kun kan overleve hvis de producerer høje koncentrationer af biotin. På denne måde kan vi finde gode producenter fra samlinger af mangfoldige stammer. Til sidst overveje vi problemet med falske positiver og sammensætter et konstrukt so kan sigte snydeceller fra.

Introduction to the thesis

Biotin biosynthesis pathway

Biotin, also known as vitamin B7 or vitamin H, is necessary for all domains of life due to its role as a cofactor to carboxylation reactions, including the essential acetyl-CoA carboxylase (acc) of central carbon metabolism. The biosynthesis pathway in *E. coli* begins by enzyme BioC to methylate malonyl-ACP, which disguises the substrate to enter fatty acid biosynthesis for two rounds of elongation (1). The precursor pimeloyl-ACP methyl ester is esterased by BioH (2) and serves as a precursor for BioF, which in *E. coli* can also accept a -CoA version of the substrate (3). The penultimate precursor, desthiobiotin, is generate by BioA and BioD activity. The final step of the pathway is catalyzed by the BioB enzyme, which generally considered the bottleneck of the pathway. BioB is a SAM enzyme which utilizes radical chemistry to activate an inert C-C bond for sulfur insertion (4). The activity of the enzyme is further characterized by two types of iron-sulfuer clusters: $[4Fe\text{-}4S]^{2+}$ is involved in the radical chemistry and $[2Fe\text{-}2S]^{2+}$ donates sulfur for the final product (5,6). Each catalytic conversion requires regeneration, aided by ferredoxins or flavodoxins (7). The BioB crystal structure indicates a dimer with two active sites (6). Experimental data indicates one active site is more catalytically active than the other; which may be due to regeneration processes (4). Further, feedback inhibition of SAM cleavage products has been observed (8), which may contribute to the low turnover rates (4).

Regulation of biotin biosynthesis occurs via the BPL (biotin protein ligase) enzyme, which serves two functions in *E. coli* where it is known as BirA (9). After binding to free biotin in the cell, it biotinylates the accB subunits. When no peptides this modification, the biotinyl-BirA self-dimerised and binds to the *bioO* operator, which represses the promoters for both halves of the biotin operon *bioA* and *bioBFCD* (10).

Previous work on biotin production in E. coli

Overexpression of *in vivo* BioB is toxic to the *E. coli* host (Chapter 3). In a separate work, a mutant strain with improved Fe-S cluster supply was identified that reduced BioB expressed and enabled higher titers. Therefore we are working under two possible hypotheses on why BioB is toxic to the cell: i) overexpression depletes intracellular pools of iron-sulfur clusters which impedes other essential

processes, ii) incorrectly folded or non-regenerated BioB may release iron into the cell which causes the formation of radical oxygen species via the Fenton reaction (11).

Biotin as a target for industrial biotechnology

The biotin molecule has been of special interest for industrial biotechnology. For one, human beings do not have the capacity to synthesise biotin. Due to the essentiality of the biotin molecule, the prokaryotic biosynthesis genes may be new antimicrobial drug targets (12–15). Secondly, in its role as a vitamin and unique co-factor, biotin and its pathway have a number of potential uses. Enzymes for polyketide synthesis in *E. coli* require non-native substrates whose heterologous pathways require biotinylation (16). Further biotin levels play an important role in fermentation. In yeast, biotin prototrophy is desirable to increase financial feasibility of processes (17). Engineering of the biotin pathway has also been beneficial for the production of odd-carbon acids in *E. coli* (18,19), streptavidin in *E. coli* (20), L-glutamic acid by *Coryneform* spp. (21), and hydrocarbon production in alga *Botryococcus braunii KMITL 5* (22). Finally, the biotin molecule itself has been the target of extensive research for strain engineering (23–28).

In **Chapter 1** of this dissertation, we review recent developments in selection systems with a particular focus on complementing traditional strain engineering with synthetic biology tools. We also highlight the difficulties of developing new selection systems. Finally we discuss the problem of false positives and selection evasion and how some recently published works overcome this challenge.
In **Chapter 2**, the native regulation of biotin biosynthesis in *E. coli* serves as a foundation for a biotin-responsive biosensor. This sensor is optimized with tools of synthetic biology to generate a robust fluorescence output. To construct a high-throughput screening methodology, the sensor is optimized for use in nanoliter reactors, which are small alginate beads that can be used for strain cultivation. In a proof-of-principle experiment, the highest producers of a mock library is enriched using a bead sorter.
Capitalising on the growth-essentiality of the molecule, we construct two selection system for biotin in **Chapter 3.** In both cases, the cellular demand for biotin is increased be decreasing the efficiency of an aspect of biotin metabolism. This allows us to identify a high precursor plasmid as well as alternatives to the native BioB enzyme. For confidentiality reasons, this chapter is undisclosed.
Finally, **Chapter 4** addresses a central problem identified in Chapter 1: selection evasion. We construct a counter-screen for a phenotype we call “tuning” where inducible BioB plasmids become inert to the inducer IPTG. In a selection scenario, the counter-screen reduces false positives 3-fold.

BIBLIOGRAPHY

[1] Lin S, Hanson RE, Cronan JE. **Biotin synthesis begins by hijacking the fatty acid synthetic pathway**. *Nature chemical biology* 2010; **6(9):**682–8.

[2] Cao X, Zhu L, Hu Z, Cronan JE. **Expression and activity of the BioH esterase of biotin synthesis is independent of genome context**. *Scientific Reports* 2017; **7(1):**1–12.

[3] Manandhar M, Cronan JE**. A canonical biotin synthesis enzyme, 8-amino-7- oxononanoate synthase (BioF), utilizes different acyl chain donors in *Bacillus subtilis* and *Escherichia coli*.** *Applied and Environmental Microbiology* 2018; **84(1):** AEM-02084.

[4] Farrar CE, Siu KKW, Howell PL, Jarrett JT. **Biotin synthase exhibits burst kinetics and multiple turnovers in the absence of inhibition by**

products and product-related biomolecules. *Biochemistry* 2010; **49(46):**9985–96.

[5] Fugate CJ, Jarrett JT. **Biotin synthase: Insights into radical-mediated carbon-sulfur bond formation.** *Biochimica et Biophysica Acta - Proteins and Proteomics* 2012; **1824(11):**1213–22.

[6] Berkovitch F, Nicolet Y, Wan JT, Jarrett JT, Drennan CL**. Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme.** Science 2004; **303(5654):**76–9.

[7] Reyda MR, Fugate CJ, Jarrett JT**. A complex between biotin synthase and the iron-sulfur cluster assembly chaperone HscA that enhances in vivo cluster assembly.** *Biochemistry* 2009; **48(45):**10782–92.

[8] Farrar CE, Jarrett JT**. Protein residues that control the reaction trajectory in s-adenosylmethionine radical enzymes: Mutagenesis of asparagine 153 and aspartate 155 in escherichia coli biotin synthase.** *Biochemistry* 2009; **48(11):**2448–58.

[9] Wang J, Beckett D. **A conserved regulatory mechanism in bifunctional biotin protein ligases.** *Protein Science* 2017; **26(8):**1564–73.

[10] Chakravartty V, Cronan JE. **The wing of a winged helix-turn-helix transcription factor organizes the active site of BirA, a bifunctional repressor/ligase**. *The Journal of Biological Chemistry* 2013; **288(50):**36029–39.

[11] Imlay JA. **Cellular defenses against superoxide and hydrogen peroxide.** *Annual Review of Biochemistry* 2008; **77(1):**755–76.

[12] Bond TEH, Sorenson AE, Schaeffer PM. **Functional characterisation of *Burkholderia pseudomallei* biotin protein ligase: A toolkit for anti-melioidosis drug development.** *Microbiological Research* 2017; **199:**40–8.

[13] Feng J, Paparella A, Booker G, Polyak S, Abell A. **Biotin protein ligase is a target for new antibacterials**. *Antibiotics* 2016; **5(3):**26.

[14] Paparella A, Soares da Costa T, Yap M, Tieu W, Wilce M, Booker G, Abell A,

Polyak S. **Structure guided design of biotin protein ligase inhibitors for antibiotic siscovery.** *Current Topics in Medicinal Chemistry* 2013; **14(1):**4–20.

[15] Liu F, Dawadi S, Maize KM, Dai R, Park SW, Schnappinger D, Finzel BC, Aldrich CC. **Structure-based optimization of pyridoxal 5′-phosphate-dependent transaminase enzyme (BioA) inhibitors that target biotin biosynthesis in *Mycobacterium tuberculosis*.** *Journal of Medicinal Chemistry* 2017; **60(13):** 5507–20.

[16] Vandova GA, O'Brien R V., Lowry B, Robbins TF, Fischer CR, Davis RW, Khosla C, Harvey CJ, Hillenmeyer ME. **Heterologous expression of diverse propionyl-CoA carboxylases affects polyketide production in *Escherichia coli*.** *Journal of Antibiotics* 2017; **70(7):**859–63.

[17] Bracher JM, de Hulster E, Koster CC, van den Broek M, Daran JMG, van Maris AJA, Pronk JT. **Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations.** *Applied and Environmental Microbiology* 2017; **83(16):** AEM-00892.

[18] Haushalter RW, Phelan RM, Hoh KM, Su C, Wang G, Baidoo EEK, Keasling JD. **Production of odd-carbon dicarboxylic acids in *Escherichia coli* using an engineered biotin-fatty acid biosynthetic pathway**. *Journal of the American Chemical Society* 2017; **139(13):**4615–8.

[19] Foster AB, Cartman ST, Kennedy J. Invista North America Sarl, 2018. **Materials and methods utilizing biotin producing mutant hosts for the production of 7-carbon chemicals.** U.S. Patent Application 2018; 15/658,958.

[20] Jeschek M, Bahls MO, Schneider V, Marlière P, Ward TR, Panke S. **Biotin-independent strains of *Escherichia coli* for enhanced streptavidin production.** *Metabolic Engineering* 2017; **40:**33–40.

[21] Niaz B, Rajoka MI, Al-Ghanim KA, Yousaf S, Mahboob S, Nadeem S. **Optimizing the concentration of biotin for L-glutamic acid production by a locally isolated coryneform strain.** *Journal of Animal and Plant Sciences* 2017; **27(4):**1217–24.

[22] Ruangsomboon S, Sornchai P, Prachom N. **Enhanced hydrocarbon production and improved biodiesel qualities of *Botryococcus braunii KMITL 5* by vitamins thiamine, biotin and cobalamin supplementation.** *Algal Research* 2018; **29:**159–69.

[23] Van Arsdell SW, Perkins JB, Yocum RR, Luan L, Howitt CL, Chatterjee NP, Pero JG. **Removing a bottleneck in the *Bacillus subtilis* biotin pathway: bioA utilizes lysine rather than S-adenosylmethionine as the amino donor in the KAPA-to-DAPA reaction.** *Biotechnology and bioengineering* 2005; **91(1):**75–83.

[24] Bower SG, Perkins JB, Yocum RR, Pero JG. DSM IP Assets BV, 2005. **Biotin biosynthesis in *Bacillus subtilis***. U.S. Patent 6,841,366.

[25] Sakurai N, Imai Y, Masuda M, Komatsubara S, Tosa T. **Construction of a biotin-overproducing strain of S*erratia marcescens*.** *Applied and Environmental Microbiology* 1993; **59(9):**2857–63.

[26] Komatsubara S, Imai Y, Masuda M, Sakurai N, Tanabe Seiyaku Co Ltd, 1994. **Microorganism and process for preparing D-biotin using the same.** U.S. Patent 5,374,554.

[27] Campbell JW, Cheung A, Eddy CK. BASF SE, 1995**. Method to produce biotin**. U.S. Patent 5,445,952.

[28] Kanzaki N, Kawamoto T, Matsui J, Nakahama K, Ifuku O. Shiseido Co Ltd and Takeda Pharmaceutical Co Ltd, 2001. **Microorganism resistant to threonine analogue and production of biotin.** U.S. Patent 6,284,500.

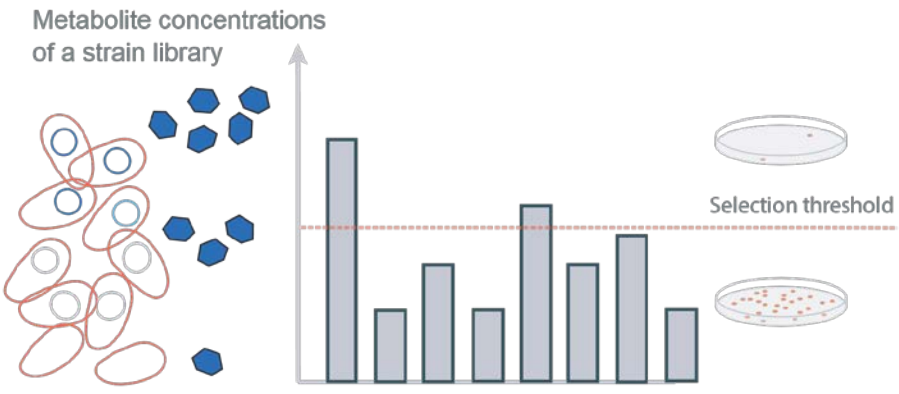
CHAPTER ONE – WIRING CELL GROWTH TO PRODUCT FORMATION

Josi Buerger, Luisa Gronenberg, Hans Jasper Genee, and Morten Sommer

This chapter has been submitted as a manuscript for Current Opinion in Biotechnology.

ABSTRACT

Microbial cell factories offer new and sustainable production routes for high-value chemicals. However, identification of high producers within a library of clones remains a challenge. When product formation is coupled to growth, millions of metabolic variants can be effectively interrogated by growth selection, dramatically increasing the throughput of strain evaluation. While growth-coupled selections for cell factories have a long history of success based on metabolite auxotrophies and toxic antimetabolites, such methods are generally restricted to molecules native to their host metabolism. New synthetic biology tools offer the opportunity to rewire cellular metabolism to depend on specific and non-native products for growth.



Graphical abstract

A diverse library of producers is characterized for production variants. Suitable selection systems allow growth only above the threshold, reducing the original library to the top producers.

INTRODUCTION

Toxic waste-products and dependency on petrochemicals are notably problematic in the chemical synthesis industry. A forward-thinking solution is replacing manufacture of materials, medicines, and biochemicals by engineered microbial processes. Despite promising benefits and a growing interest, widespread industrial implementation of biotechnology has been constrained by long development timelines and process economics due to the challenges of strain engineering (1).

The expansion of the DNA modification toolbox allows for genetic editing in a targeted, rational manner and this has generated a profusion of engineering strategies based on innovation in DNA synthesis and genome editing (2). Yet the technology to investigate the resulting clones is often based on low-throughput, traditional analytics which cannot address the number of clones present in such diverse libraries.

Biological reporter systems that link product formation to an immediately detectable output represent a potential solution to accelerate laborious screening protocols using either fluorescence-activated cell sorting (3) or microfluidic droplet based sorting (4). However, the reporter systems with the largest throughput are those that allow for growth selection by wiring product formation to cell growth (5, 6). With a focus mainly on *Escherichia coli*, we describe recent work on applications of conventional selection strategies, review synthetic selection systems, and illustrate the difficulties that arise from the construction and use of such approaches when deployed for industrial purposes.

Expanding on conventional selection systems

Conventional selection systems have been employed in industrial biotechnology since the 1970's and often rely on auxotrophic knock-outs and inhibitory molecules to construct dependencies on the target molecule. One implementation of gene deletions is the discovery of novel sequences for glycerol utilisation using functional metagenomic selections (7,8) (Figure 1A).

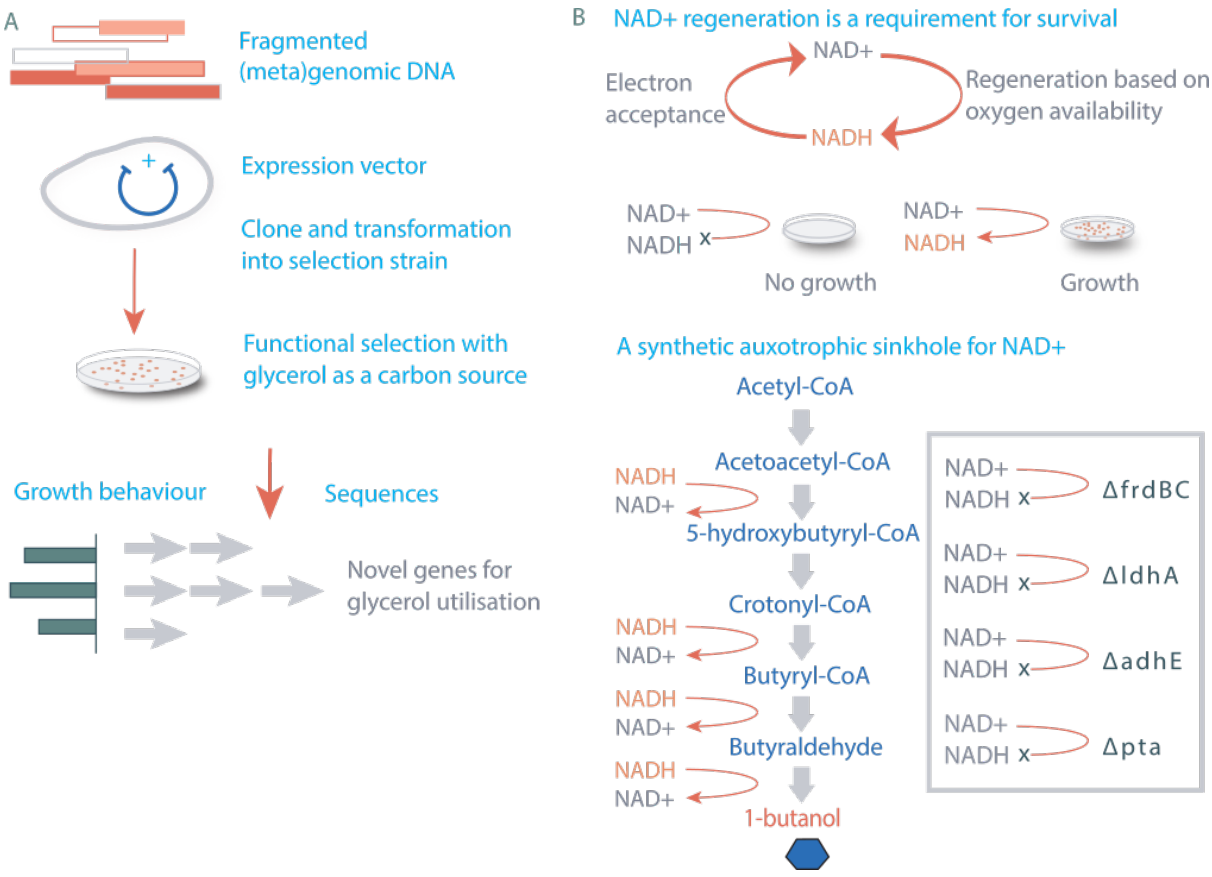


Figure 1: Utilising auxotrophy for growth-dependent selection

A - Functional metagenomics selections can be utilized in metabolic engineering. Here, DNA fragments of a metagenomic library are selected for improved capacity to utilize glycerol as a carbon source (9)(8).

B - The NAD+/NADH redox requirement of E. coli is utilised to build a strain dependent on 1-butanol product formation. As other mixed-fermentation reactions are knocked out (grey box), NAD+ can only be regenerated by the 1-butanol pathway (10).

Auxotrophic strains are a powerful tool for selection, but not all molecules of interest are amenable for this type of growth-coupling. Hence, the construction of non-native auxotrophies has been pursued. For example, 1-butanol is an important fermentation product from *Clostridia spp.* The native pathway is strongly CoA-dependent and this co-factor demand has challenged optimization or heterologous transfer. Particularly in *E. coli,* the metabolic state of high CoA is only available during anaerobic recycling of NADH via mixed-acid fermentation. In several studies, multiple knockouts of electron acceptors leave 1-butanol as the only “NADH outlet” for redox balance under anaerobic growth (10,11). Without high flux through the 1-butanol pathway, lack of NAD+ regeneration leads to growth arrest. The redox dependency enabled the selection of mutants

with increased activity of CoA reductase from error-prone PCR mutagenized libraries and fermentation titers for butanol were improved to 88% of the theoretical yield, reaching 30 g/L. The authors note the suitability of this anaerobic growth rescue approach for other NADH-dependent products, including lactate, alanine, or succinate.
The success of an industrial process can hinge on the choice of a suitable production strain and overlap of target molecules with metabolic dependency should be considered. Computational resources that identify relevant gene knock-outs are useful tools (12,13). Predictive genome-scale modelling, particularly constraint-based reconstruction and analysis (COBRA) methods, can probe further options for engineering a synthetic metabolic link for the compound of interest (14–17). These tools are introducing powerful computational resources for a technology traditionally based in the lab.

Another selection approach to strain engineering relies on antimetabolites, which are metabolite analogues that inhibit growth (18). This can be due to incorrect substrate recognition leading to enzymatic inhibition or disruption of pathway regulation. At the correct concentration, antimetabolites force the cell to elevate enzyme concentrations or small-molecule products to overcome the inhibition burden. For example, three structural antimetabolites enabled selection of strain variants with improved production of aspartic acid production (19). In another study, riboflavin production was increased to 680 mg/L in *Candida famata* by the use of structural analogues of riboflavin in combination with a colour-based screen for the vitamin (20). Antimetabolites can also be used to select for increased tolerance to toxic pathway intermediates (21, 22).

Another type of selection can be imposed by light-excitable quantum cadmium telluride dots. The dots generate ROS-stress with the assumption that increased tolerance to superoxide is correlated with strong NADPH metabolism, key for industrial molecules. A loss-of-function mutation in *hdfR* was identified using this selection and exhibited a 2-fold increase in titers for 3-hydroxypropionic acid, a NADPH-limited pathway (23).

Conventional selections are powerful tools of metabolic engineering and expanding them with synthetic biology tools is pushing their versatility even further. Muconic acid production in *S. cerevisiae* was improved three-fold by

combining a biosensor conferring geneticin resistance with ALE against 4-fluorophenylalanine, a competitive inhibitor of aromatic amino acids. The evolved strains gave titers of 2.1 g/L (21). Next-generation sequencing gives insight into population behaviour (19) and selections can be combined with rational engineering (19, 20). Nevertheless, important limitations remain for both antimetabolite and auxotrophic selection. Auxotrophic demands exist only for essential metabolites and this molecular requirement must lie in the range of titers relevant to industrial processes. Unfortunately, this is rarely the case and accordingly selection options arising from auxotrophic or inhibitory elements can be limited.

Synthetic coupling of growth to product formation

Synthetic growth-coupled systems have the same aim as conventional strain engineering: identification of high producers from a pool of clones. Starting from ligand-responsive switches, gene networks are rewired to monitor the presence of target molecules. Gene switches can be based on riboswitches (Figure 2) or transcription factors (TFs) (Figure 3), which are then linked to selectable marker genes to generate a synthetic or non-natural coupling of growth to product formation.

A seminal selection system utilised the specific response of the NahR TF to benzoic acids but not the corresponding aldehydes and coupled it to tetracycline resistance (24). The selection identified increased enzymatic activity of *xylC*, a benzaldehyde dehydrogenase from *P. putida*. A separate system focused on lysine production repurposed a riboswitch upstream of *lysC* to build a "Riboselector" such that the presence of lysine reduces toxic expression of *tetA* (25). In four rounds, a proof-of-concept plasmid expressing varying promoter strengths for *ppc*, a key node for the lysine pathway, was enriched. In the same work, a tryptophan-specific Riboselector was generated from an aptamer functioning as an ON switch.

A selection system capable of single cell selection was developed for thiamine based on the natural ThiM19 riboswitch (6,26,27). In the presence of thiamine pyrophosphate, the repurposed ThiM19 riboswitch allows for translation initiation

of an otherwise repressed gene. In the study, the switch was coupled to an antibiotic resistance cassette, thereby generating a synthetic selection strain growth-dependent on thiamine pyrophosphate. To apprehend selection escapees, the system was expanded to a second thiamine pyrophosphate riboswitch with a different antibiotic resistance gene, which reduced the rate of selection evasion by 1,000-fold. The established selection system was expanded for xanthine alkaloid selection (6) and applied to metagenomic gene discovery and transport engineering (28); demonstrating its versatility and selection power.

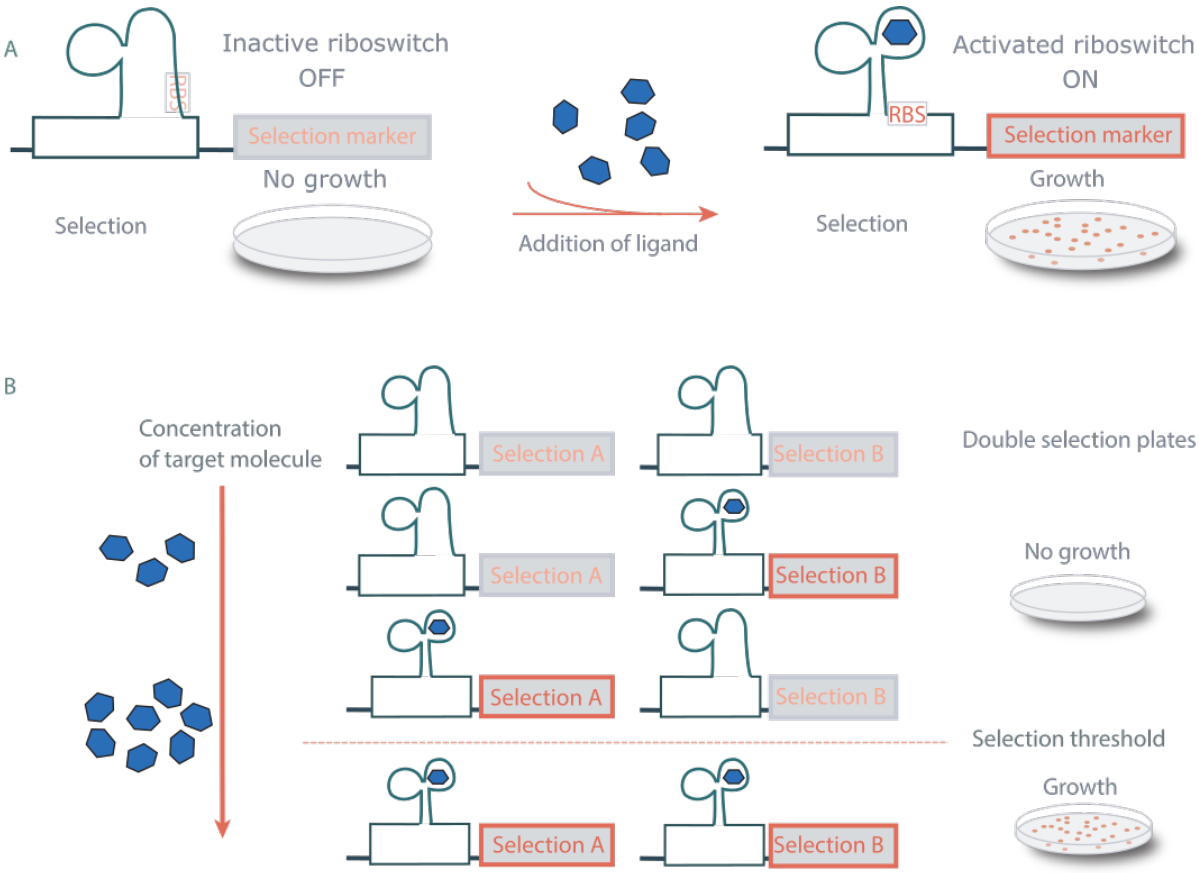


Figure 2: Riboswitch-based selection

A - A riboswitch is a regulatory element that acts at the level of RNA. Addition of the molecule of interest facilitates a structural change of the transcript, which can uncover (in this example) an RBS binding site. Shown (left) is a riboswitch without ligand in its OFF state so that the selection marker is not active (light grey). The addition of ligand allows expression of the selection marker, switching the selection to the ON state (dark orange) and allowing growth on selection plates.

B - In order to generate a robust selection platform that avoids a high rate of false positives, Genee et al. (6) place two separate antibiotic resistance cassettes under control of independent thiamine pyrophosphate riboswitches, shown to control Selection A and Selection B. When no product is present, both riboswitches are inactive. In the

case of point mutations, a single riboswitch would lose its selection pressure and allow growth on selection plates. Instead, the set-up confers resistance to both selection antibiotics only for i) high product concentrations and ii) no loss of selection to activate both riboswitches to the ON state.

As with the NahR example, synthetic selections often utilize TF-based metabolite sensing to conditionally express antibiotic cassettes (29–32). A heterologous activator and promoter from *Thauera sp.* conditionally expressed *tetA* in the presence of 1-butanol. This system was used to screen Δ*adhE* strains with plasmid-encoded RBS libraries of heterologous *kivD* and *ADH6* from yeast. Increased activity of 35% was identified (30). Two additional systems were built to showcase the utility of conditional *tetA* expression. Firstly, the native *E. coli* system for succinate is regulated by two-component DcuR/DcuS which was coupled to *tetA* expression. Secondly, a system for adipate was built from *Pseudomonas putida* PcaR (30).

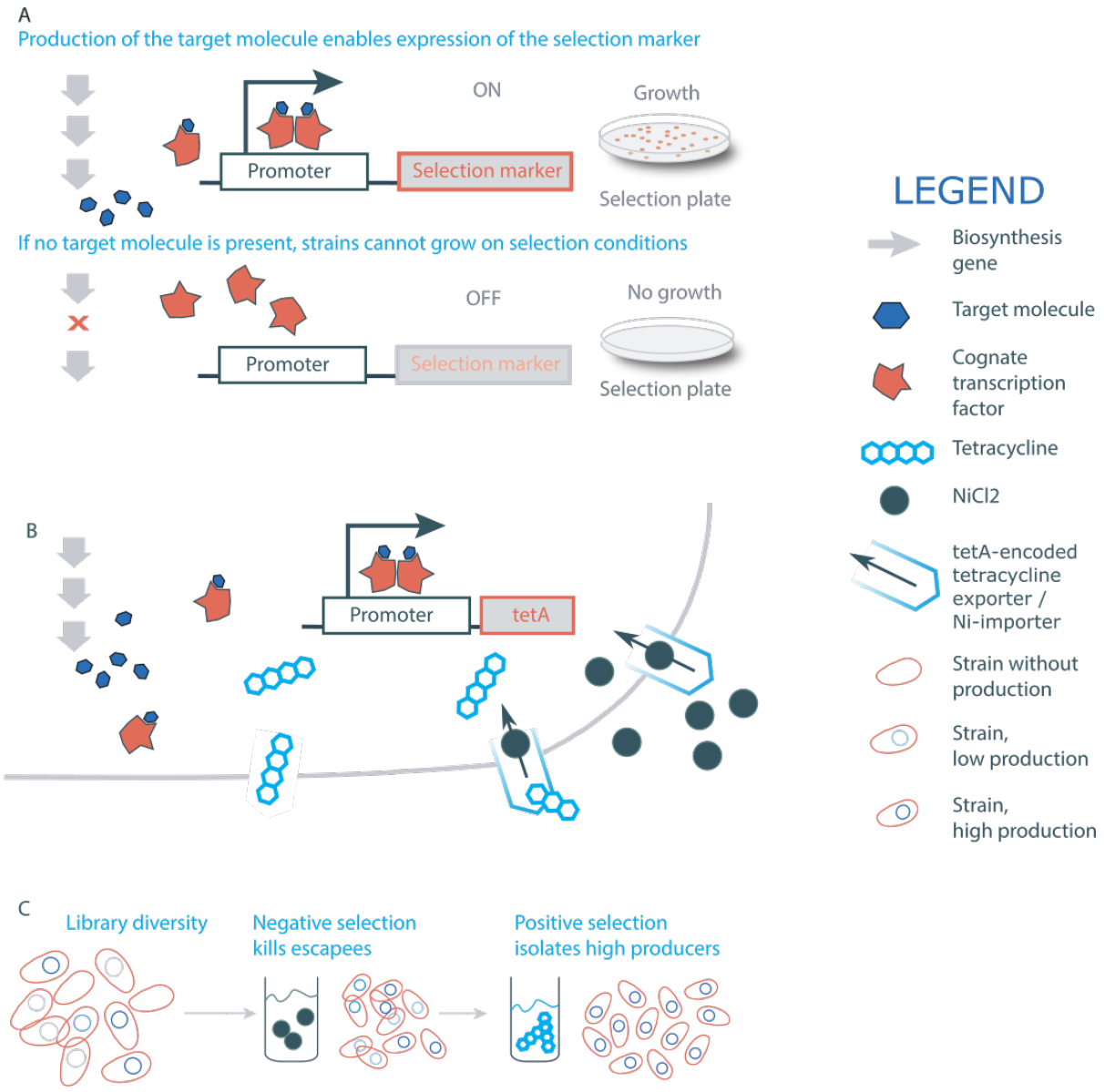


Figure 3: Transcription-factor based selection

A - Biosynthesis of a target molecule is coupled to a cognate transcription factor which often dimerises to activate or repress a selection marker. Shown is a positive selection marker where the presence of the target molecule enables expression of an antibiotic resistance marker, allowing only producers to grow on a selection plate.

B - Negative selection strategies arise from markers with dual effects on the cell. For example, gene tetA provides resistance to tetracycline but makes the host cell susceptible to $NiCl_2$ toxicity. Negative selection functions ensure that selection "winners" have not bypassed the selection pressure.

C - A workflow for a joint selection system with a positive and negative round is shown. From a diverse library, high producers are isolated using a round of positive selection. Prior to screening, a negative selection step ensures that the selection pressure has been maintained to rule out false positives.

Nevertheless, evolutionary escapees can overwhelm selection systems (6,31). A "toggled" approach reduced this issue by alternating between positive and negative selection rounds (Figure 3C). In addition to enriching for high producers, the selection kills the fraction of non-producers that have mutated to evade the selection system (33). Using both *tolC* and kanamycin resistance under a single conditional module can generate a kanamycin-resistant strain susceptible to colicin-E1 (31). Selection winners can be tested for false positivity by presence of colicin-E1 susceptibility. Toggled selection was used to identify a 36-fold naringenin producer from genome-engineered libraries using *tolC* under regulation of TtgR. In a second example, *tolC* regulated by CdaR increased glucaric acid 22-fold (31).

Synthetic selections have also been applied for increased protein expression and export. To avoid extraction protocols, secretion is advantageous for recombinant protein production. A common approach is to fuse the protein of interest to naturally secreted substances such as YebF or OsmY, but this suffers from low yields (34). Coupling of YebF to BLiP, a β-lactamase inhibitor protein, conveys increased resistance against β-lactam antibiotics. This approach was used to find mutants with improved extracellular accumulation of desired proteins in libraries with 10^{12} members (34).

Development of future synthetic selection systems

At the heart of many selection systems are genetic switches harvested from nature. Yet, the construction of novel selection systems from an identified switch is challenging. Nevertheless, recent developments point towards efficient design strategies that could be applied to expand the repertoire of selections.

Despite the prevalence of regulatory elements across diverse genomes, the identification or construction of robust metabolic sensors can be laborious for either protein or RNA based sensors (35). SIGEX (substrate-induced gene-expression screening), which places a promoterless reporter gene on either side of a randomized metagenomic region in the presence of the desired target, is a powerful tool. If a transcription-factor-like sequence is present, the reporter gene will be expressed (36). The identification of specific aptamers using SELEX

(systematic evolution of ligands by exponential enrichment) and the rational design of riboswitches is also not trivial (37,38), as the relationship between sequence and function of RNA is not readily discerned (39). Modelling RNA sequence to function may assist in the future construction of riboswitch-based selection systems by tapping into the power of algorithmic approaches (40,41). An approach called “term-seq” identifies natural ribo-regulators in genome sequences via early termination events. This methodology is independent of evolutionary conservation and can find riboswitches that comparative genomics would not (37,38).

After identification, biosensors can still behave unexpectedly in the production host. Insufficient specificity can reduce the applicability of a biosensor. For example, the adipate TF-sensor also responded to pimelate (30). Nevertheless, non-specific biosensors can be powerful for some applications. One study used a proxy biosensor for a pathway intermediate linked to expression of geneticin resistance to identify overproduction strains (21). Further, co-cultivation strategies can overcome the difficulties of having selection and production in a single strain. A ultra high-throughput application of fluorescence markers was used to screen multiplexed libraries and identify a *B. subtilis* producer for riboflavin via alginate co-encapsulation with an *E. coli* sensor strain (42). Proxy biosensors and co-cultivation may offer solutions when appropriate TFs or riboswitches are not available.

A separate issue concerns biosensor range, *i.e.* metabolite concentrations which lead to discernable signal changes (Figure 4A). Dynamic ranges are rarely appropriate for continuous selection systems with increasing titers. A fluorescent biosensor for L-valine production in *Corynebacterium glutamicum* enabled increased titers by 25% for 5 rounds of FACS sorting before the upper detection limit had been reached (43). The 1-butanol sensor described above (30) could not be used above 25 mM metabolite concentration. The glucaric acid system (31) also reports the range’s limitations. A proffered solution is transport regulation to direct the rate of molecule uptake and thereby the dynamic sensor range (31) (Figure 4B). A further strategy is the construction of molecular buffer systems. In this work (44) a TF is split into two functional domains, DNA-binding or activation. An excess of DNA binding domain mimics chemical pH buffer systems to generate a robust signal and protection against promoter leakiness.

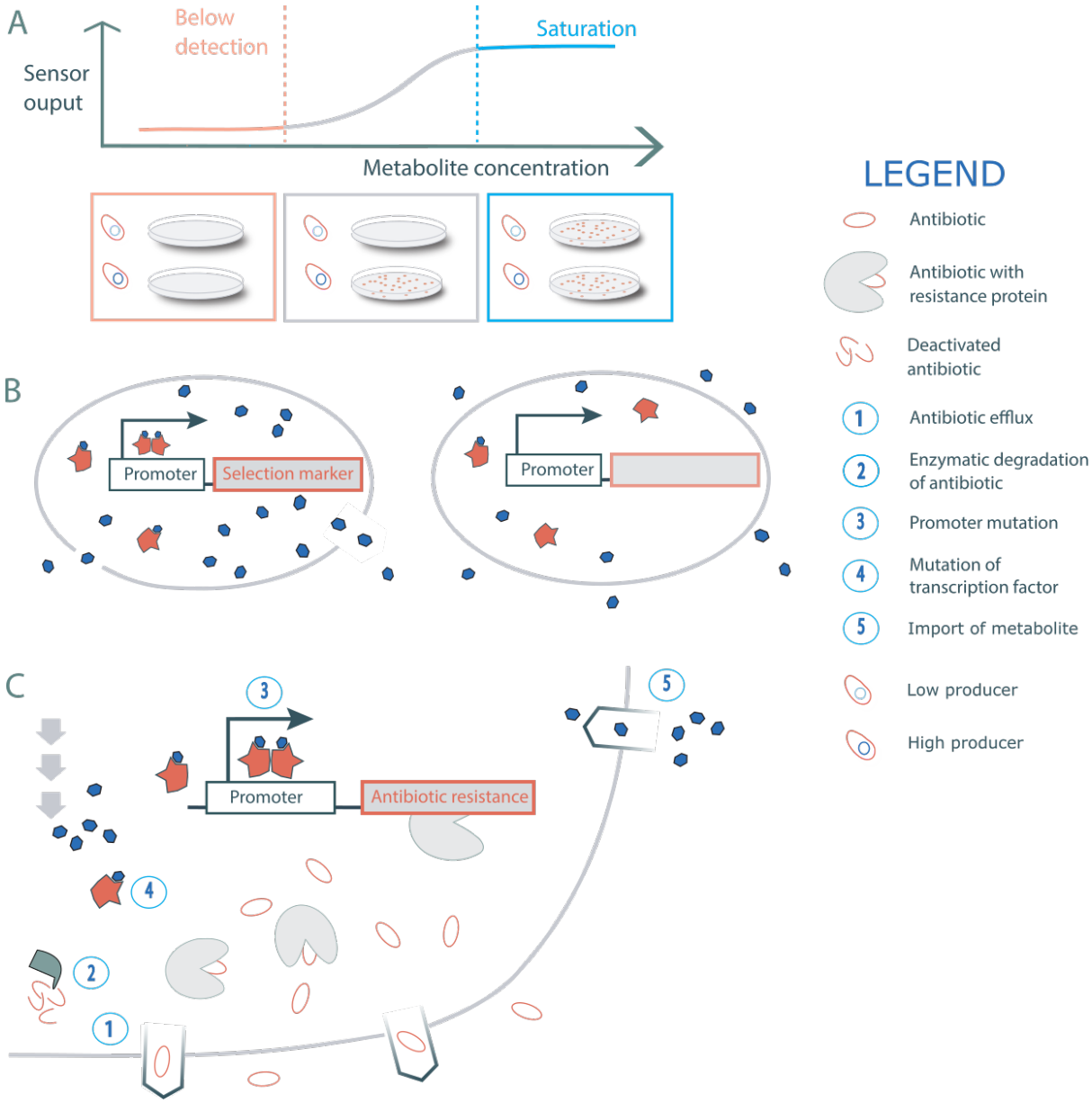


Figure 4: Tuning selection systems and limiting evolutionary escape

A - The graph indicates output of a selection marker based on metabolite concentration. Two saturation points are indicated; a lower threshold where decrease in production does not lead to decrease of output (orange dotted line) and an upper threshold where increase in production does not lead to increase in output (blue dotted line). The coloured boxed below show the effect of high or low production at each of the three phases, emphasising the importance of biosensor range to get a screenable output correlated with product formation.

B - As in Raman et al. (31), engineering of transporters can shift the functional range of a selection system. Two sensors are shown, with (left) or without (right) active import, which

changes the conditional expression of the selection marker even though the extracellular metabolite concentrations remain the same.

C - Loss of selection via evolutionary escapees

Overview of a cell with a generic growth-couples selection system: a transcription factor binds to the molecule of interest, dimerises, and activates the expression of an antibiotic resistance cassette. Numberings 1-5 illustrate potential mechanisms of selection escapee. These mechanisms would allow the cell to survive an antibiotic challenge even if it did not produce the molecule of interest.

Finally, a major obstacle is the intrinsic response of biological systems to overcome selective pressure (45,46). A single point mutation in promoter regions can confer loss of selection, which provides the escapee with unrivalled growth advantage compared to the remaining population (Figure 4C). The thiamine riboswitch gave a false positive rate of 10E-3 for an early, unoptimised selection construct (6). In nature, genetic networks often display redundancy so that their function is robust to mutagenic effects. Gene duplication is a strong opposition to genetic drift, and this strategy can be utilised to reduce selection escape. Redundancy within selections can consist of two antibiotic resistance cassettes under one regulator (6) or single genes with two controllable outputs (33) . Raman *et al*. (31) demonstrate the versatility of selection doubling by expanding ten different biosensors in a modular way. The concept of redundancy is a powerful addition to the toolkit of selection systems.

CONCLUSION

Within industrial biotechnology, the need for robust, modular, and usable selection systems arise with the ability to manipulate, design, and multiplex the DNA code. Selection strategies of conventional strain engineering can identify genotypes with higher production capacity and are amenable for expansion with modern techniques. The further development of synthetic circuits to wire growth to product formation allows for the characterization of millions of genetic variants, with particular applications to generate sustainable cell factories for a greener chemical industry. Nevertheless, the success of synthetic selection systems will rely on overcoming the challenges of their construction and application: identification of relevant conditional modules, solving issues of specificity and range of *in vivo* biosensors, and addressing the false positives and negatives that arise from loss of selection.

BIBLIOGRAPHY

[1] Nielsen J, Keasling JD. **Engineering Cellular Metabolism.** *Cell* 2016; **164(6):**1185–97.

[2] Lennen RM, Nilsson Wallin AI, Pedersen M, Bonde M, Luo H, Herrgård MJ, Sommer MOAS. **Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects.** *Nucleic Acids Research.* 2015; **(11):**e36.

[3] Lehning CE, Siedler S, Ellabaan MMH, Sommer MOA. **Assessing glycolytic flux alterations resulting from genetic perturbations in *E. coli* using a biosensor.** *Metabolic Engineering* 2017; **42**:194–202.

[4] Siedler S, Khatri NK, Zsohár A, Kjærbølling I, Vogt M, Hammar P, Nielsen C, Marienhagen J, Sommer MOAS, Joensson H. **Development of a bacterial biosensor for rapid screening of yeast p-coumaric acid production.** *ACS Synthetic Biology* 2017; **6(10):**1860–9.

[5] Rogers JK, Church GM. **Genetically encoded sensors enable real-time observation of metabolite production.** *Proceedings of the National Academy of Sciences* 2016; **1;113(9):**2388–93.

[6] Genee HJ, Bali AP, Petersen SD, Siedler S, Bonde MT, Gronenberg LS, Kristensen M, Harrisonn SJ, Sommer MOAS. **Functional mining of transporters using synthetic selections**. *Nature Chemical Biology* 2016; **12:**1015–22.

[7] Brady SF, Clardy J. **Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA**. *Angewandte Chemie* 2005; **44(43):**7063–5.

[8] Helm E Van Der, Genee HJ, Sommer MOA. **The evolving interface between synthetic biology and functional metagenomics.** *Nature Chemical Biology* 2018; **14**: 752–759.

[9] Loaces I, Rodríguez C, Amarelle V, Fabiano E, Noya F. **Improved glycerol to ethanol conversion by *E. coli* using a metagenomic fragment isolated from an anaerobic reactor.** *Journal of Industrial Microbiology and Biotechnology* 2016; **43(10):**1405–16.

[10] Shen CR, Lan EI, Dekishima Y, Baez A, Cho KM, Liao JC. **Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli.*** *Applied and Environmental Microbiology* 2011; **77(9):**2905–15.

[11] Wen RC, Shen CR. **Self-regulated 1-butanol production in *Escherichia coli* based on the endogenous fermentative control.** *Biotechnology for Biofuels* 2016; **9(1):**1–15.

[12] Burgard AP, Pharkya P, Maranas CD. **OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization.** *Biotechnology and Bioengineering* 2003; **84(6):**647–57.

[13] Hassanpour N, Ullah E, Yousofshahi M, Nair NU, Hassoun S. **Selection Finder (SelFi): A computational metabolic engineering tool**

to enable directed evolution of enzymes. *Metabolic Engineering Communications* 2017; **4**:37–47.

[14] Feist AM, Zielinski DC, Orth JD, Schellenberger J, Markus J, Palsson BØ. **Model-driven evalution of the production potential for growth coupled products of *Escherichia coli.*** *Metabolic engineering* 2011; **12(3):**173–86.

[15] Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE … & Palsson BØ. **Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0.** *Nature Protocols* 2011; **6(9):**1290–307.

[16] Brunk E, George KW, Alonso-Gutierrez J, Thompson M, Baidoo E, Wang G, Petzold CJ … Keasling JD, Palsson BØ & Lee TS. **Characterizing strain variation in engineered *E. coli* using amulti-omics-based workflow.** *Cell Systems* 2016; **2(5):**335–46.

[17] Ma D, Yang L, Fleming RMT, Thiele I, Palsson BØ, Saunders MA. **Reliable and efficient solution of genome-scale models of metabolism and macromolecular expression**. *Scientific Reports* 2017; **7:**40863.

[18] Fiedurek J, Trytek M, Szczodrak J. **Strain improvement of industrially important microorganisms based on resistance to toxic metabolites and abiotic stress**. *Journal of basic microbiology* 2017, **57.6:** 445-459.

[19] Bonomo J, Lynch MD, Warnecke T, Price J V., Gill RT. **Genome-scale analysis of anti-metabolite directed strain engineering.** *Metabolic Engineering* 2008; **10(2):**109–20.

[20] Dmytruk K V., Yatsyshyn VY, Sybirna NO, Fedorovych D V., Sibirny AA. **Metabolic engineering and classic selection of the yeast Candida famata (*Candida flareri*) for construction of strains with enhanced riboflavin production.** *Metabolic Engineering* 2011; **13(1):**82–8.

[21] Leavitt JM, Wagner JM, Tu CC, Tong A, Liu Y, Alper HS. **Biosensor-enabled directed evolution to improve muconic acid production in *Saccharomyces cerevisiae.*** *Biotechnology Journal* 2017; **12(10):**1–9.

[22] Commichau FM, Alzinger A, Sande R, Bretzel W, Reuß DR, Dormeyer M, Chevreux B, Schuldes J, Daniel R, Prágaia Z. **Engineering *Bacillus subtilis* for the conversion of the antimetabolite 4-hydroxy-l-threonine to pyridoxine.** *Metabolic Engineering* 2015; **29:**196–207

[23] Reynolds TS, Courtney CM, Erickson KE, Wolfe LM, Chatterjee A, Nagpal P, Gill RT. **ROS mediated selection for increased NADPH availability in *Escherichia coli.*** *Biotechnology and Bioengineering* 2017; **114(11):**2685–9.

[24] van Sint Fiet S, van Beilen JB, Witholt B. **Selection of biocatalysts for chemical synthesis.** *Proceedings of the National Academy of Sciences* 2006; **103(6):**1693–8.

[25] Yang J, Seo SW, Jang S, Shin S-I, Lim CH, Roh T-Y, Jung GY. **Synthetic RNA devices to expedite the evolution of metabolite-producing microbes**. *Nature communications* 2013; **4:**1413.

[26] Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. **Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch.** *Nature* 2006; **441(7097):**1167–71.

[27] Muranaka N, Sharma V, Nomura Y, Yokobayashi Y. **An efficient platform for genetic selection and screening of gene switches in *Escherichia coli.*** *Nucleic acids research* 2009; **37(5):**e39.

[28] Bali AP, Genee HJ, Sommer M. **Directed evolution of membrane transport using synthetic selections**. *ACS Synthetic Biology* 2018; **7(3):**789-793.

[29] Dietrich JA, McKee AE, Keasling JD. **High-throughput metabolic engineering: advances in small-molecule screening and selection**. *Annual Review of Biochemistry* 2010; **79:** 563-590.

[30] Dietrich JA, Shis DL, Alikhani A, Keasling JD. **Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis.** *ACS Synthetic Biology* 2013; **2(1):**47–58.

[31] Raman S, Rogers JK, Taylor ND, Church GM. **Evolution-guided optimization of biosynthetic pathways.** *Proceedings of the National Academy of Sciences* 2014; **111(50):**201409523.

[32] Rogers JK, Guzman CD, Taylor ND, Raman S, Anderson K, Church GM. **Synthetic biosensors for precise gene control and real-time monitoring of metabolites.** Nucleic Acids Research. 2015;43(15):7648–60.

[33] Gregg CJ, Lajoie MJ, Napolitano MG, Mosberg J a., Goodman DB, Aach J, Isaacs FJ, Church GM. **Rational optimization of tolC as a powerful dual selectable marker for genome engineering**. *Nucleic Acids Research* 2014; **42(7):**4779–90.

[34] Natarajan A, Haitjema CH, Lee R, Boock JT, DeLisa MP. **An engineered survival-selection assay for extracellular protein expression uncovers hypersecretory phenotypes in *Escherichia coli***. *ACS Synthetic Biology* 2017; **6(5):**875–83.

[35] de Paepe B, Peters G, Coussement P, Maertens J, de Mey M. **Tailor-made transcriptional biosensors for optimizing microbial cell factories.** *Journal of Industrial Microbiology and Biotechnology* 2016; **44(4):**1–23.

[36] Uchiyama T, Watanabe K**. Substrate-induced gene expression (SIGEX) screening of metagenome libraries.** *Nature Protocols* 2008; **3(7):**1202–12.

[37] Dar D, Shamir M, Mellin JR, Koutero M, Stern-Ginossar N, Cossart P, Sorek R. **Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria.** *Science* 2016; **352(6282):** aad9822.

[38] Sommer MOA, Suess B**. (Meta-)genome mining for new ribo-regulators.** *Science*; **352(6282):** 144-145.

[39] Schneider C, Suess B. **Identification of RNA aptamers with riboswitching properties.** *Methods* 2016; **97**: 44-50.

[40] Adjeroh D, Allaga M, Tan J, Lin J, Jiang Y, Abbasi A, Zhou X. **Feature-based and string-based models for predicting RNA-protein interaction**. *Molecules* 2018; **23(3):**1–17.

[41] Choi D, Park B, Chae H, Lee W, Han K. **Predicting protein-binding regions in RNA using nucleotide profiles and compositions.** *BMC Systems Biology* 2017; **11:**1–12.

[42] Meyer A, Pellaux R, Potot S, Becker K, Hohmann H-P, Panke S, Held M. **Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors.** *Nature Chemistry* 2015; **7(8):**673–8.

[43] Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J, Frunzke J. **Biosensor-driven adaptive laboratory evolution of l-valine production in *Corynebacterium glutamicum.*** *Metabolic Engineering* 2015; **32:**184–94.

[44] Rugbjerg P, Genee HJ, Jensen K, Sarup-Lytzen K, Sommer MOA. **Molecular buffers permit sensitivity tuning and inversion of riboswitch signals.** *ACS Synthetic Biology* 2016; **5(7):**632-638.

[45] Ryu YS, Chandran SP, Kim K, Lee SK. **Oligo- and dsDNA-mediated genome editing using a tetA dual selection system in *Escherichia coli*.** *PLoS ONE* 2017; **12(7):**1–16.

[46] Drake JW. **A constant rate of spontaneous mutation in DNA-based microbes**. *Proceedings of the National Academy of Sciences* 1991; **88(16):**7160–4.

CHAPTER TWO – SCREENING IN NANOLITER REACTORS WITH A BIOTIN BIOSENSOR

This chapter is printed in its original form.

Abstract

Biological reporter systems that link product formation to an immediately detectable output represent an acceleration to laborious screening protocols. With some modification, biosensors are amenable for high-throughput screening using alginate beads (nanoliter reactors) to characterise the production capacity of thousands instead of hundreds of clonal variants. Here we present the development of a biotin-sensitive biosensor, optimisation of a nanoliter reactor protocol, and proof-of-concept screening of a library.

INTRODUCTION

Transcription-factors as stepping stones to biosensors

Biosensors are a versatile tool to identify high producers from a library of clonal variants (1). At their heart, biosensors have a domain to detect the metabolite of interest and a conditionally expressed, readily discernible output. Output domains are often well-characterised genetic markers such as assayable *lacZ* (2) or fluorescent proteins, e.g. GFP from jellyfish *Aequorea victoria* (3). Detection domains often rely on repurposing gene regulation, as many genetic events are coupled to metabolite sensing. In particular, transcription factors and their paired promoters are key for the construction of novel biosensors. For example, a transcription factor PadR from *B. subtilis* for *p*-coumaric acid was expressed in *E. coli* to control the expression of a *yfp* to generate measurable fluorescence only in the presence of *p*-coumaric acid (4).

To act as a robust biosensor amenable for nanoliter reactors, a sensing-output domain must fulfill a number of requirements. The first requirement is orthogonality and suitability with respect to host metabolism. Orthogonality is the capacity for seamless recombinant expression. Transcription factor - promoter pairs should be transferable to multiple hosts without change in functionality. The transfer should not interact with or disrupt native metabolism or regulation. This is of particular relevance for metabolic engineering: the presence of a biosensor should not encumber the rational restructuring towards the molecule of interest. The choice between native and recombinant sensing-output domains must be deliberated. Heterologous constructs expand the range of sensible molecules, but the effect of synthetic expression can be problematic (reviewed e.g. 5) However, native transcription factors often act as global regulators within host metabolism so that their overexpression impacts the global transcriptome (5). In situations where a native TF does *not* affect regulation, the native concentration may be insufficient for signal mediation (20). Not only the effect of the sensing domain has to be taken into consideration: fluorescent proteins can also impact bacterial physiology, *e.g.* mCherry can give false OD measurements (33).

Table 1: *In vivo* biosensors using TF in *E coli* published in the last 5 years

Target molecule	Genetic architecture	Application
N-acetylneuraminic acid	Engineered promoters that interact with native NanR to express mKate2 (5)	Identify high N -acetylneuraminic acid producers
Acrylate	AcuR from *Rhodobacter sphaeroides* and sfGFP (6)	
Arsenite	Native ArS repressor and operator control *egfp* (7)	Arsenite
Cadmium	CadC homologue from *Bacillus oceanisediminis 2691* and GFP (8)	Heavy metal sensing in microfluidics application
Cellubiose	CelR from *Thermobifida fusca* with GFP (1)	Biosensor for detection of cellulolytic activity
Clorpyrifos (CPF)	Cognate promoter chpA and activator ChpR from *Sinorhizobium meliloti* with fluorescent output 4-MU activated by a native sulferase (9)	Pesticide detection in environmental samples
p-coumaric acid	PadR from *B. subtilis* activates yfp expression (4)	
3,4-dihydroxy benzoate	pcaU and promoter from *Acinetobacter sp.* ADP1 with GFP (10)	FACS to improve sensor
Doxycycline	TetH/TetR modules from *Haemophilus somnus 2336* and GFP (11)	Selective detection of tetracycline variants
Erythromycin	Native MphR and sfGFP (6)	
Formaldehyde	Native FrmR and error-prone PCR library of formaldehyde responsive promoters with GFP (12)	FACS to improve sensor
Fructose-1,6-bisphosphate	Cra activator and *ppsa* promoter with GFP (13)	Monitor effect of different carbon sources or mevalonate production on glycolytic flux
Glucarate	Native CdaR and sfGFP (6)	Monitor flux in glucarate production
Heme	Heme-binding domain cytochrome b562 from *Saccharomyces cerevisiae* linked to EGFP and mKATE2 (14)	Study of heme metabolism
3-hydroxypropionate	Native prpR, pprpR, and pcs from *Chloroflexus aurantiacus* or a version with acuR from Rhodobacter sphaeroides (15)	Identify optimal conditions for 3-HP production
Itaconic acid	*Yp*ItcR/P_{ccl} from *Yersinia pseudotuberculosis* with RFP (16)	Identify suitable gene expression for biosynthesis of itaconic acid in *E. coli*
Lactams (ε-caprolactam, δ-valerolactam, butyrolactam)	Activator-promoter pair ChnR/P_B from *Acinetobacter sp.* strain NCIMB9871 (17)	
Lead	CadC homologue from *Bacillus oceanisediminis 2691* and GFP (8)	Heavy metal sensing in microfluidics application
Macrolides	Native MphR responsive to *e.g.* erythromycin, josamycin, oleandomycin,	Error-prone PCR of the activator binding domain combined with

	narbomycin, methymycin and pikromycin with GFP (18)	FACS generates increased activator specificity
Maltose	Synthetic fusion of DNA-binding zinc finger module with protein for maltose binding (19)	Maltose
Monosodium methylarsonic acid (MSMA)	Sequence-modified ArsR of *Acidithiobacillus ferrooxidans* and according operator controlling GFP expressed (20)	In vitro biosensor for herbicide detection
NADPH	SoxR and *eyfp* (21)	Screen a mutant enzyme library for improved activity
Naringenin	FdeR from *Herbaspirillum seropedicae* SmR1 and fluorescent marker (22)	Identify suitable biosynthesis genes for flavonoid production
Naringenin	TtgR from *Pseudomonas putida* and sfGFP (6)	
Nickel	RcnR C35A and rcnA promoter with *lux* reporter gene in *E. coli* TD2158 (23)	Lower limit of detection for nickel detection in drinking water
p-nitrophenol	Mutated PobR from *Acinetobacter* with GFP (24)	Characterise hydrolysis activity on insecticide
2-oxaloacetate	2-oxaloacetate binding domain found in NifA from *Azotobacter vinelandii* with *yfp* (25)	Study effect of nutritional changes on metabolite pool in *E. coli*
Phenol	DmpR activator with GFP (26)	Evolve novel tyrosine-phenol-lyases
Phenylalanine	TroR activator with *mtr* promoter activates Venus (27)	Screen promoter library for biosynthesis of phenylalanine
Phenylalanine	TyrR activator and *tyr* promoter with YFP (28)	Screening of a mutagenized library with RBS variants for production in *E. coli*
Qercetin and kaempferol	QdoR from *Bacillus subtilis* and fluorescent marker (22)	FACS of kaempferol producers
Shikimic acid	Mutated HucR and promoter from *Deinococcus radiodurans* expressed RFP (29)	Improve flux into shikimic acid pathway in *E. coli* via engineered AroG
Triacetic acid lactones	Various AraC mutants activate GFP (30) and further engineering (31)	FACS to diversify substrate recognition and carry out metabolic engineering for TALs

A second requirement is suitable dynamic range and response for the desired application. The dynamic range of a sensor is the concentration of metabolite which generates a detectable difference in output, also called the transfer function (6). The relevant concentration will vary between intracellular and extracellular applications depending on molecular transport (34). For example, a promiscuous biosensor can require different amounts of two flavonoids for the minimal activation concentration: in fact a 2-fold increase between kaempferol and quercetin (22). Further, a study with identical sensor plasmid architecture but

different heterologous sensing domains showed that the optimal measurement time ranged from 1 - 6h (6); indicating that time is a relevant factor for range.

Thirdly, a sensor must show the appropriate target specificity. Conditional output expression should occur only in the presence of the metabolite of interest. Adequate specificity is pressing concern for many applications: for example, to reliably detect herbicides using a biosensor, detection needs to be restricted to the biologically toxic trivalent arsenic species (MSMA and roxarsone) but not inorganic arsenic (20). Hence, biosensors are often tested for inertness to structurally related compounds: for example, a cellubiose biosensor was tested for response to the structurally similar glucose polymers and disaccharides (1).

Finally, the inherent susceptibility of biological systems to genetic drift and mutation pressures should not be disregarded, in particular if a sensor will be used *in vivo* over extended cultivation time. This has been already discussed in Chapter 1.

Applications - nanoliter reactors

Once a robust, orthogonal biosensor has been developed, it may still require further optimisation for its desired application, often the interrogation of large clonal libraries. Although library screening with a sensor can be done via strain cultivation and a fluorescence plate-reader, throughput of this approach is limited to 10^3 per day for manual workflows and 10^4 for automated protocols. A solution is the use of droplet or micro-bead technology. Sorting with nanoliter reactors can interrogate up to 10^6 events per day, an incomparable advantage for screening large libraries.

Nanoliter reactors or alginate beads[1] are formed from an anionic polymer isolated from brown algae. They form a viscous substance that can be mixed with biological matter to generate stable globules by passing through an extrusion nozzle, separation by directed air-flow, and hardening in $CaCl_2$ buffer (35). This generates individual, separated reactors that can isolate proteins, chemical or enzymatic reactions, proteins, and entire bacterial cells. When coupled with a measurable output and a sorting device, encapsulation are an excellent tool for

[1] Also referred to as hydrogels or microgels

screening large libraries with wide applicability. Various whey proteins were analysed in the beads and tested for stability (36), a laccase from the white-rot basidiomycete *Coriolopsis gallica* was identified for superior dye degradation ability, and the GFP variant sfGFP was optimised for acidic conditions (37). Even structurally sensitive RNA aptamers have been successfully encapsulation and sorted in nanoliter beads (38).

Whole-species encapsulation include *Pseudomonas aeruginosa, E. coli,* and *Laccococcus spp.* Whole-cell biocatalysis for target products is a powerful implementation of alginate beads, particularly for *E. coli.* For example, production of l-xylulose was achieved by expressing recombinant l-arabinitol dehydrogenase and NADH oxidase to regenerate the needed co-factors (35). Further applications include α-ketoisocaproate production (39), cadaverine production (40), and even high-throughput colony PCR (41). Within strain engineering, the use of alginate beads can be developed further beyond cell immobilisation. Alginate beads can serve as nanoliter reactors for multiple strains and species. For example, a library of *Bacillus subtilis* vitamin B12 producers was screening with an FMN-responsive riboswitch expressed in an *E. coli* strain. Using FACS-like sorting, highly fluorescent beads are isolated to provide the high production variants (38). Here we apply the co-cultivation strategy to enrich *E. coli* strains with plasmids for biotin production using an *E. coli* biosensor.

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli strain One Shot™ TOP10 (Invitrogen) was used for cloning and plasmid extraction. Biosensor characterisation was carried out in *E. coli* strain BW25113 from the *Keio* collection (1) as well as JW3803-2 ΔyigM with kanamycin cassette removed. Production strain for the cultivation was carried out in BS1575[2]. Sensor strain in the NLRs was BW25113 *29Δglk2Δpts1ΔmanZ1.*

Table 2: Plasmids constructed in Chapter 2

Plasmid	Construction
pBS195	Amplification of *sfGFP* (2) and ampicillin resistance cassette from pBS046 (oBS20, oBS21) Amplification of colE1 from pBS19 (pZA31luc from EXPRESSYS) (oBS1105, oBS1107) Amplication of *bioO-bioA* region from *E. coli* MG1655 (oBS1161, oBS1162)
pBS249	Amplification of *sfGFP* and ampicillin resistance cassette from pBS046 (oBS20, oBS21) Amplification of colE1 from pBS19 (pZA31luc) (oBS1105, oBS1007) Amplication of *bioO-bioB* region from *E. coli* MG1655 (oBS768, oBS789)
pBS334	Amplification of p15A from commercial vector pBAD18 (ThermoFisher) (oBS1172, oBS1173)
pBS411	Amplification of smR resistance cassette from pBS259 (pMA7SacB (3)) (oBS1159, oBS1160)
pBS459 - pBS464	Amplication of mCherry with library of promoters (oBS1440, oBS1435)

Chemicals

Chemicals were purchased from Carl Roth GmbH unless otherwise noted. Biotin (Sigma-Aldrich) and DTB (Sigma-Aldrich) were made as 100 mM DMSO stock solutions. Antibiotics ampicillin, kanamycin, zeocin, and spectinomycin (Sigma-Aldrich) were dissolved as 1000x stock solutions in MQ. Plating experiments were carried out by mixing 2x-concentration MOPS or LB with 3% agar (Sigma-Aldrich). All solvents were handled according to manufacturer's instructions.

[2] Strain genotypes are available in the Annex.

Media and cultivation

Molecular biology protocols were carried out in 2xYT made from 16 g tryptone, 10 g yeast extract, and 5 g NaCl per liter; or SOC media containing 20 g tryptone, 5 g yeast extract, and 0.5 g NaCL per liter, supplemented with 0.5 mL 2 M $MgCl_2$ and 1 mL 20% D-glucose per 100 mL before use. Biosensor characterization was carried out in minimal MOPS (e.g. 4) containing 10x MOPS, 10 mL 0.132 M K_2HPO_4, 500x vitamin solution, and 10 mL 20% D-glucose. Nanoliter optimization was carried out in M9 containing 10x M9 salts (Sigma Life Sciences), 20 mL 20% D-glucose, 0.1 mL 1 M $CaCl_2$, 2 mL 1M $MgSO_4$, and 0.5 mL trace element solution. CDM (chemically defined media, recipe in Annex) was also used. Stock protocols are placed in the Annex.

Molecular biology techniques

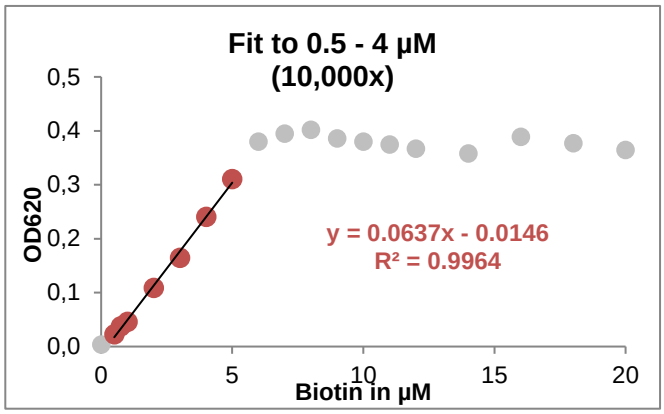
DNA fragments were generated by PCR from genomic DNA (*E. coli* MG1655 or BW25113) or plasmid DNA using primers synthesized by IDT or Eurofins for Gibson Assembly (5) or USER cloning (6). USER primers were designed with AMUSER software (7). PCR experiments for Phusion U polymerase (Thermo Fisher) contained HF-buffer (Thermo Fischer), dNTPs (Thermo Fischer), 0.02 U/µL polymerase, 0.5 µM of each primer, and template DNA at concentrations per manufacturer's instructions. PCR cycles followed PhuU amplification protocol in a Ristretto thermocycler (VWR). PCR products mixed with 6x DNA Gel Loading Dye (Thermo Fischer) were subject to gel electrophoresis in 1% agarose (Lonza) with SYBR Safe DNA Gel Stain (Thermo Fischer). Electrophoresis in TAE buffer diluted from 10x stock was carried out at 90 V, 400 mA for 35 minutes. Gels were visualized using blue light and 1 kB GeneRuler DNA ladder (Thermo Fisher), excised if necessary, purified with appropriate Cycle Pure Kits (Omega), and concentrations determined *via* Nanodrop (MySpec VWR). Plasmid assembly was carried out according to USER (6) or Gibson assembly (5) manuals. Strains were prepared following protocols for chemical competency (8) or electroporation (9) and transformed with ligated or plasmid DNA. Plasmid extractions were performed with E. Z. N. A. Plasmid DNA Mini Kit (Omega). Sanger sequenced was performed with Mix2seq kits (Eurofins).

Characterisation of growth and fluorescence

Strain characterization was carried out 96-well lidded deep-well plates (Enzyscreen). Colonies were inoculated into precultures of the respective media and incubated at 37 °C for 18-20 h at 274 RPM (Innova Thermo Fischer). OD was determined at a 1:10 dilution using a plate reader (Multiskan FC, Thermo Scientific)[3]. Production cultures were inoculated at a concentration 1:100, normalised by preculture OD where required. Strain fluorescence was determined with a plate reader (VarioSkan LUX, Thermo Scientific) at emission/excitation at 485/518 nm or 575/618 nm for sfGFP and mCherry, respectively. Kinetic data for growth and fluorescence were obtained by cultivation in microtiter plates (greiner bio-one) with porous seals (BreatheEasy; Diversified Biotech). Measurement was carried out every 10 minutes.

Bioassay

To assay biotin concentrations in supernatant samples, growth of a Δ*bioB* strain with zeocin resistance to extracellular biotin at various dilutions was used. Biotin standards between 0.5-30 µM and supernatant samples were added to each master plate, which was typically diluted and inoculated 1:10 with the bioassay strain at a final dilution of 7,500 – 25,000. As a correlation between growth and biotin availability for *E. coli* is obtainable between ~50-500 pM, multiple dilutions are necessary. Examples below shows growth response to dilution range at 10,000x.



[3] Values were converted to cuvette OD in all cases using previously determined conversion rate: CuvetteOD = MULTISKAN/0.2569 - 0.0001

Encapsulation, extraction, and analysis of nanoliter reactors

For encapsulation at respective λ = 300 and λ = 0.3, glycerol stocks of sensor and production strains were diluted in final volume of 2 mL sterile 0.85% NaCl, vortexed, and added to 8 g 2.5% alginate (Sigma Aldrich). Using a plastic syringe, mixture was pumped through a laminar jet-break encapsulator (Nisco Engineering AG) at flow rate 2 ml/min, mode P9, and a 0.35 mm nozzle into stirred 100 mM CaCl2 for hardening. After 15 minutes, beads were washed with 10 mM $CaCl_2$ and soaked in respective media. Excess media was thoroughly washed off and beads allowed to dry prior to incubation in liquid (2mL 10 mM CaCl2) or hydrophobic (50 mL heavy mineral oil with 2% ABIL 90 and 0.05% Tween 20) phase. Encapsulations were incubated at 37 °C 200 RPM. For imaging and COPAS sorting, beads were extracted from hydrophobic phase *via* centrifugation at 80g for 10 minutes, then washed with 10 mM $CaCl_2$. Fluorescence assay and sorting was carried out with a COPAS (Union Biometrica). Identified beads were sorted into 96-well deep-well plates containing minimal media with glucose which does not allow sensor strains to grow.

RESULTS

The native regulatory elements of *E. coli* biotin biosynthesis are amenable to biosensor construction

The native operon structure of biotin biosynthesis in *E. coli* is separated into three transcription events at two genomic loci (42). The *bioABFCD* region is strictly regulated due to the high synthesis cost of the biotin molecule. We utilise BirA, the repressor protein of the operon promoters to build a biotin-responsive biosensor. To ensure functionality of the operator site, output gene *sfGFP*[4] (43) is placed after the first ten amino acids of the regulated gene. Given the bidirectional nature of the *bioO* operator (Figure 1), two promoter regions P_{BA} and P_{BB} were compared for signal mediation (Figure 2, V1). The response of V1 constructs to biotin concentrations was poor, with large variability between replicates (CV>50%) and small difference between *on* and *off* states. Comparison of the promoter options (Linear fit of 24 data points, R2=0.52 compared to R2=0.28 of PBA) indicated P_{BB} as a candidate for further optimisation.

It was hypothesised that differences in plasmid copy number between individual cells could account for noise in GFP output (e.g. 20). A common normalisation approach is to constitutively express a second fluorescent marker, in this case mCherry. The sequence was added under the control of constitutive promoters from the Anderson library (45) at three different expression strengths (BBa_J23105 with 623 AU, BBa_J23106 with 396 AU, BBa_J23117 with 162 AU). At the lowest expression level, mCherry was inert to biotin concentrations (data not shown) and used to build V2. The GFP output is normalised by overall expression capacity such that the subsequent transfer function of the sensor is given by the GFP:mCherry (Figure 2, V2).

[4] Superfolder GFP

Figure 1: Native biotin regulation is engineered to build sensor plasmid pBS453

Left - Transcription of biotin biosynthesis is tightly regulated in *E. coli.* In conditions of excess biotin (shown in yellow), transcription factor BirA binds the metabolite and self-dimerises (complex shown in darker grey) to bind and repress the *bioO* operator encompassing promoters for *bioA* and the *bioBFCD* operon. When the *bioO* domain is free, transcription occurs (yellow arrows).

Right - Basic plasmid template of V3 with p15A origin of replication and spectinomyin resistance cassette *smR* is used to express *bioO*, the operon responsive to dimerised BirA-biotin complexes. The sfGFP output is fused to the first 10 amino acids of the BioB sequence to ensure orthogonal transfer of the repressive region. *mCherry* is added to the plasmid under a medium level promoter to normalise for plasmid copy number and the relaying transcription factor *birA* is added in the same operon. Resulting plasmid pBS453 has a total size of 4982 bp.

To address the noise of the V2 construct in the off state (concentrations 10^1-10^6 µM), we considered the importance of the BirA transcription factor for signal mediation. Each *bioO* repression event requires two biotin molecules and two BirA proteins for dimerization and the copy number of p15A origin of replication lies at ~20 as well as the *bioO* on the genome. Predicted cellular levels of BirA as a non-abundant transcription factor are <50, which led us to hypothesise that at high biotin levels, BirA concentrations may be limiting to relay the signal. The BirA sequence was added to the V2 construct by placing it in an operon with the mCherry gene with a DNA spacer[5]. An RBS library of 6 members was tested at predicted TIR[6]s 19, 95, 308, 337, 949, and 1255 (46). The addition of the transcription factor did not shift the range of the sensor, but more stable on and off states were observed for predicted RBS strength of >300 (V3, right). At a lower RBS, the quality of the GFP signal was in fact reduced (V3, left) Finally, an option to increase the range of the sensor was tested by transferring the plasmid into a knockout of the *E. coli* biotin importer *yigM*. This shifted the range of the

[5] gtttggttgaattc
[6] Translation initiation rate

sensor 100-fold (Figure 2, bottom): both sensors are in the off state at 1000 µM; the *yigM* WT is on at 100 µM whereas *yigM* KO is on at 0.1 µM extracellular biotin. Further, the relative difference between the on and off fold states compared to V1 was increased 100-fold (7 to 70 normalised fluorescence units).

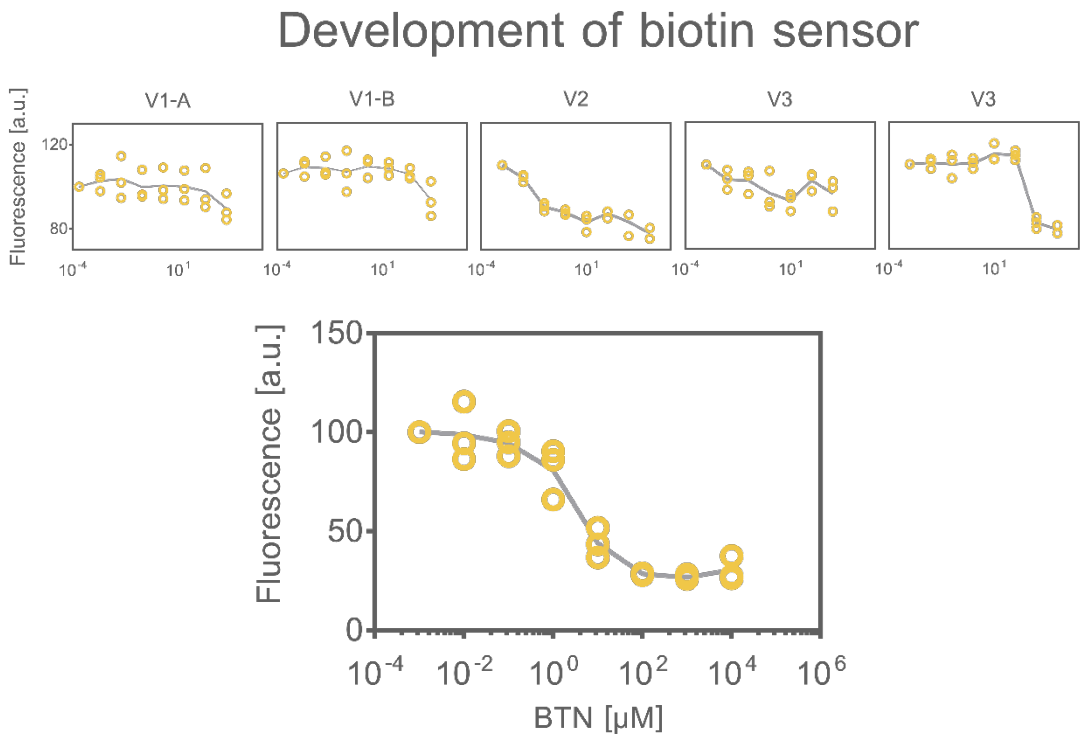


Figure 2: Stages of development for a biotin-responsive biosensor

Top row - A plasmid-based system containing native *E. coli* biotin operator regulation expression of GFP gave an average signal (grey line) only weakly correlated to extracellular biotin concentration with low reproducibility across replicates (yellow data points, V1-A, B). Improvements to the sensor are shown by inclusion of mCherry (V2) and transcription factor BirA (V3, efect of two RBS strengths is shown). Fluorescence data is normalised to *off* state; axes are identical.

Bottom graph - The optimised sensor construct generates a clear GFP output (n=3) to biotin concentrations and improves the signal-to-noise ratio as well as expanding the range of output.

Co-cultivation in nanoliter reactors

Factors for co-cultivation protocol

In collaboration with FGen GmbH we developed our biotin biosensor for high-throughput sorting in nanoliter reactors. An analogous biotin sensor was built by FGen with increased sensitivity, ampicillin resistance, and reverted output with a tetR repressor module (range 0-50 nM in *yigM* WT, 0-250 nM in *yigM* KO[7]). The modified sensor plasmid sAS278 was expressed in a glucose inert strain BW25113 *29Δglk2Δpts1ΔmanZ1*. Genetic engineering of carbon utilisation allowed for co-cultivation of the sensor strain with a second *E. coli* production strain. The biotin production strain was grown on glucose (0.001%) while the sensor grew on any other usable carbon source supplied, in this case 1% galactose. Within defined minimal media (M9 or CDM), both producer and sensor strain grew, albeit to different colony size (Figure 3). Co-cultivation for sorting demands a single producer colony in each bead filled with ~300 sensor beads. Given the occupation distribution of strains in beads, average occupation λ = 0.3 ensured minimal doubly occupied beads. As shown in Figure 3C, sensor fluorescence positively responded to extracellular biotin.

[7] Range determination is alginate beads is heavily dependent on experimental conditions.

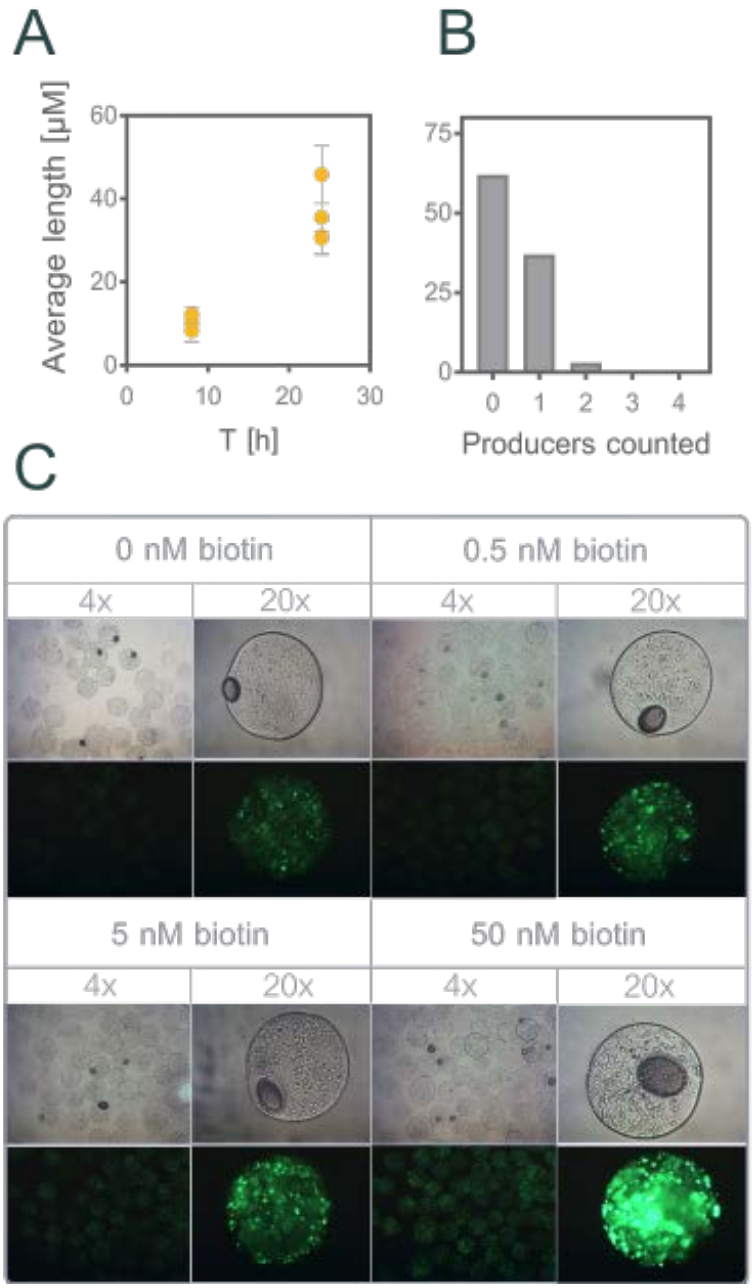


Figure 3: Media and protocol optimisation for co-cultivation

A - Average colony diameter is determined at 8h and 24h after encapsulation. For each sample, 5 colonies in standard beads with single occupation were measured (x and y axis). To correct for potential subjectivity of this measurement process, the furthest outlier from the average was deleted from each measurement time-point. Outliers identified by: largest of [|(AVERAGE(x-measurement; y-measurement))-AVERAGE(all data points from that time-point)|]

B - Bead occupation of a producer strain is determined by counting colonies in a total of 102 beads (work carried out.

C- Brightfield microscopy and fluorescence images of beads co-cultivating a biotin-responsive sensor, *E. coli* strain BW25113, and extracellular biotin. Beads were incubated in hydrophobic p ha se for 48h and extracted for imaging.

Library for proof-of-concept screening

To showcase the suitability of a TF biotin sensor in nanoliter reactors, we assembled a mock library for biotin production. The *E. coli* biosynthesis genes were expressed on a low copy-number plasmid with the same antibiotic resistance as the sensor plasmid (*ampR*). Diversity of the library is due to two alternative operon structures, native arrangement of *E. coli* (CU plasmids) or a synthetic operon structure (DU plasmids). In each case, the *bioH* sequence is placed after an RBS library to balance with the remaining pathway. The bottleneck BioB sequence is also expressed separately in each case. Plasmids were transformed into production strain BW25113 *Δbio**.

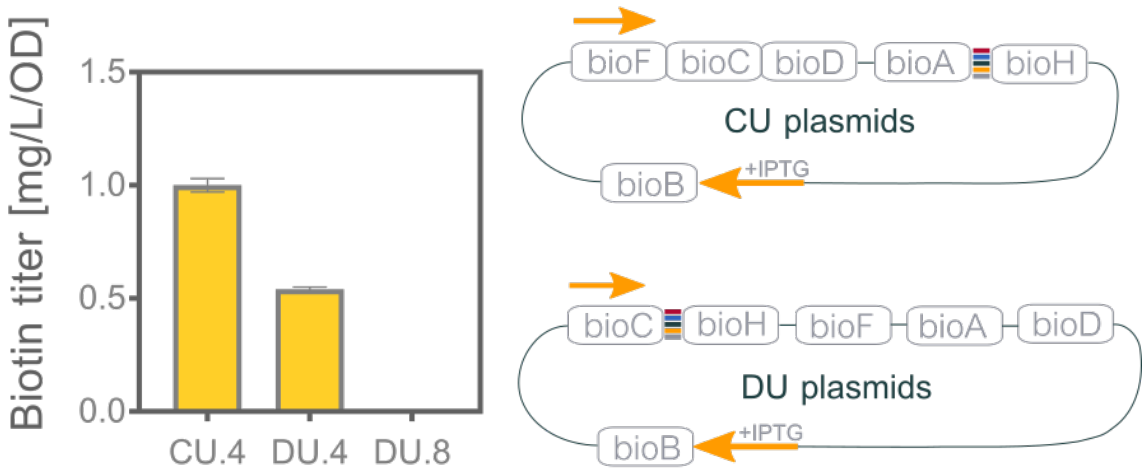


Figure 4: Proof-of-concept library

Left - Biotin titer for three representative library members. Titers (n=4) were determined using the bioassay protocol.

Right - Plasmid architecture of combinatorial library used in the co-cultivation set-up. Biotin biosynthesis genes are expressed in native (CU) or synthetic (DU) structures with bioH placed after an RBS library. In each case, *bioB* is expressed separately under IPTG induction. Titer of DU.8 was below detection.

To determine suitability of the mock library for encapsulation, growth behaviour was monitored by measuring colony diameter of each library member and the background strain (Figure 3A). No statistical differences in colony size was observed at 8h or 24h. The larger error bars observed after 24h are attributable to high levels of variability in the measurement protocol. Overall the library was deemed appropriate for a proof-of-principle sorting experiment.

COPAS sorting and verification

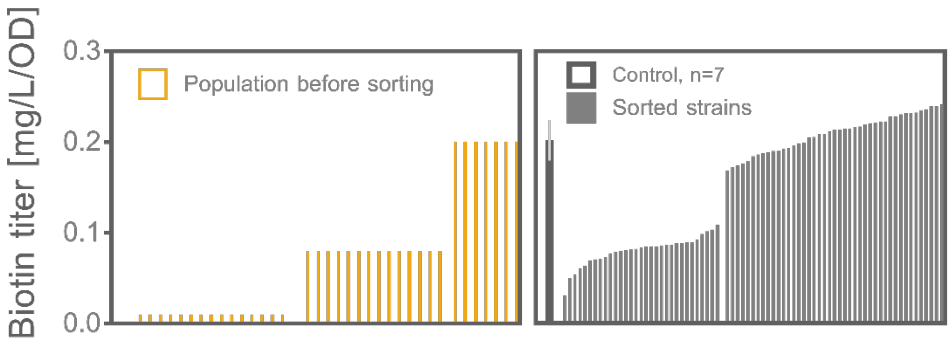


Figure 5: Library before and after COPAS sorting

Left – The distribution of strains carrying plasmids DU.4 (low), DU.8 (medium), and CU.4 (high) before encapsulation. Note that the schematic is not based on actual data and serves to visualize the enrichment.

Right - Production values of 70 strains recovered after COPAS bead sorting and grown in minimal media. Production values grouped in two sets corresponding to the expected production values of DU.8 and CU.4. Control was CU.4 plasmid not subject to sorting. Plasmids were confirmed with sequencing, corresponding to an enrichment of the high producers from 20% to ~50%.

Sensor strain and library were co-encapsulated and incubated for 48h. Library members were mixed in a ratio of 2:2:1 (DU4:DU8:CU4) to simulate biological diversity. As a control, library members were also co-encapsulated individually. The resulting fluorescence was imaged after 48 h in hydrophobic phase. Using a COPAS flow pilot (Union Biometrica), 6000 events were quantified for fluorescence following time-of-flight size gating. Top 5% of library beads were sorted into recovery media. To isolate producer strains, beads were incubated in minimal media with glucose as the sole carbon source. Out of 76 sorted events from the top 5% gate, 74 could be recovered (OD = 2.85 ±0.22), and 70 tested for production using the bioassay (Figure 5), corresponding to sample loss <10%. As strain CU.4 produced 2-fold more than DU.8 and had a mean fluorescence of 3976 AU compared to 1080 AU for DU.8, we concluded that the production was stable both before and after sorting (compare Figure 4). After sorting, no low producers (DU.4) were recovered. For strains carrying plasmid CU.4, sorting resulted in an enrichment of 20% to ~50% ($p<0.0001$, unpaired t-test with Welch's correction). Plasmid identity was confirmed via sequencing. The proof-of-concept sorting was deemed successful.

DISCUSSION

The use of nanoliter reactors to interrogate production libraries combines the advantages of selection - efficient identification of high producers due to high-throughput - with the advantages of screening - relative ease to establish and ability to characterise the whole library. In case of sensitive enzymatic steps, immobilisation in alginate can even increase yields (39). Other options of high-throughput screening such as FACS, microfluidics, and droplet sorting, offer similar advantages. Nevertheless, establishing protocols for nanoliter reactors is a non-trivial step of successful screening. Depending on the demands of the screen, factors such as pH (36), bead size (41), ratio of alginate to sample (47), and temperature (40) should be considered.

Challenges of encapsulating in nanoliter reactors

In this Chapter, we detailed the construction of a novel transcription factor-based biosensor for our target molecule. Although utilising native gene expression events is common to build sensor domains, we obtained only a weak GFP response to biotin concentrations in initial construct V1. The signal was amplified and stabilised by the combination of synthetic biology tools and overexpression of the transcription factor. This emphasised the relevance of TFs for fine-tuned signal mediation in orthogonal sensors (20). It should be addressed that a well-functioning TF-based biosensor is not immediately amenable for alginate bead application. In our case, additional modifications were necessary to transfer the sensor functionality. Generally, the required sensor range for a bead can be entirely different to a liquid cell culture. The ratio of sensor occupation to producers in a bead must be considered. Secondly, media demands for bead stability and engineering of central carbon metabolism of the sensor will also affect the transfer function. Co-cultivation approaches in particular require sugar orthogonality, which can affect growth, production, and sensing behaviour of both strains. Thirdly, stability of production can impact the screening especially if target molecules induce toxicity at high concentrations (6). Utilising KO strains can help to maintain production plasmids in case of instability (12). Finally it should be noted that the system detailed here, as opposed to previously published two-species co-cultivation (38), is a co-cultivation of two *E. coli* strains with its own particular challenges.

Overall we gated and analysed 5% of the library, leading to an enrichment from 20% to 50% of the highest producer. To enrich further, more stringent sorting conditions could be considered. However due to the nature of bead occupation (Poisson distribution), approximately one third of beads will always remain unoccupied (only sensor is present). Therefore, a gate decrease to less than 2.5% is barely feasibly, as the overall sorting time increases exponentially. If a higher percentage of CU.4 is desired from the sorting, multiple enrichment rounds need to be carried out instead. The suitability of this system for high-throughput library screening will depend heavily on the characteristics of the library.

Characterising the mock library

To elucidate the genotypic cause of production differences between DU and CU-plasmids, we sequenced the plasmids after sorting. Firstly, we recognise that the native biosynthesis operon outperforms the refactored pathway, as has been previously shown for endogenous pathways (48). Further differences were identified between DU.4 and DU.8 for RBS strengths of BioC and BioH. Well-performing DU.8 had balanced expression between the two genes (predicted TIR of 32,000 and 43,000 AU respectively) whereas DU.4 showed a skewed balance towards BioC (88,000 and 16,000 AU respectively). An additional unexpected mutation in the promoter region of DU.4 may add to mis-balancing of pathway intermediates, which could explain the very low production of this strain.

In this Chapter, we have established protocols for nanoliter media, strain encapsulation, and producer extraction with a high survival rate using only sugar modification. We are confident that this workflow will simplify future screening - selection projects.

BIBLIOGRAPHY

[1] Kwon KK, Yeom SJ, Lee DH, Jeong KJ, Lee SG. **Development of a novel cellulase biosensor that detects crystalline cellulose hydrolysis using a transcriptional regulator.** *Biochemical and Biophysical Research Communications* 2018; **495(1):**1328–34.

[2] Casadaban MJ, Cohen SN. **Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli.*** *Journal of Molecular Biology* 1980; **138(2):**179–207.

[3] Patterson GH, Lippincott-Schwartz L. **A photoactivatable GFP for selective photolabeling of proteins and cells**. *Science* 2002; **297(5588):**1873–7.

[4] Siedler S, Khatri NK, Zsohár A, Kjærbølling I, Vogt M, Hammar P, Nielsen C, Marienhagen J, Sommer MOAS, Joensson H. **Development of a bacterial biosensor for rapid screening of yeast p-coumaric acid production.** *ACS Synthetic Biology* 2017; **6(10):**1860–9.

[5] Peters G, De Paepe B, De Wannemaeker L, Duchi D, Maertens J, Lammertyn J, De Mey M. **Development of N-acetylneuraminic acid responsive biosensors based on the transcriptional regulator NanR.** *Biotechnology and Bioengineering* 2018; **115(7):**1855–65.

[6] Rogers JK, Guzman CD, Taylor ND, Raman S, Anderson K, Church GM. **Synthetic biosensors for precise gene control and real-time monitoring of metabolites.** *Nucleic Acids Research* 2015; **43(15):**7648–60.

[7] Merulla D, Van Der Meer JR**. Regulatable and modulable background expression control in prokaryotic synthetic circuits by auxiliary repressor binding sites**. *ACS Synthetic Biology* 2016; **5(1):**36–45.

[8] Kim HJ, Lim JW, Jeong H, Lee SJ, Lee DW, Kim T, Lee SJ. **Development of a highly specific and sensitive cadmium and lead microbial biosensor using synthetic CadC-T7 genetic circuitry.** i 2016; **79:**701–8.

[9] Whangsuk W, Thiengmag S, Dubbs J, Mongkolsuk S, Loprasert S. **Specific detection of the pesticide chlorpyrifos by a sensitive genetic-based whole cell biosensor.** *Analytical Biochemistry* 2016; **493:**11–3.

[10] Jha RK, Kern TL, Fox DT, M. Strauss CE. **Engineering an *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry.** *Nucleic Acids Research* 2014; **42(12):**8150–60.

[11] Hong H, Park W. **TetR repressor-based bioreporters for the detection of doxycycline using *Escherichia coli* and *Acinetobacter oleivorans.*** *Applied Microbiology and Biotechnology* 2014; **98(11):**5039–50.

[12] Rohlhill J, Sandoval NR, Papoutsakis ET. **Sort-Seq approach to engineering a formaldehyde-inducible promoter for dynamically regulated *Escherichia coli* growth on methanol.** *ACS Synthetic Biology* 2017; **6(8):**1584–95.

[13] Lehning CE, Siedler S, Ellabaan MMH, Sommer MOA. **Assessing glycolytic flux alterations resulting from genetic perturbations in *E. coli* using a biosensor.** *Metabolic Engineering* 2017; **42:**194–202.

[14] Hanna DA, Harvey RM, Martinez-Guzman O, Yuan X, Chandrasekharan B, Raju G, Outten FW, Hamza I, Reddi AR. **Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors.** *Proceedings of the National Academy of Sciences* 2016; **113(27):**7539–44.

[15] Rogers JK, Church GM. **Genetically encoded sensors enable real-time observation of metabolite production**. *Proceedings of the National Academy of Sciences* 2016; **1;113(9):** 2388–93.

[16] Hanko EKR, Minton NP, Malys N. **A transcription factor-based biosensor for detection of itaconic acid.** *ACS Synthetic Biology* 2018; **7(5):**1436–46.

[17] Zhang J, Barajas JF, Burdu M, Ruegg TL, Dias B, Keasling JD.

Development of a transcription factor-based lactam biosensor. *ACS Synthetic Biology* 2017; **6(3):**439–45.

[18] Kasey CM, Zerrad M, Li Y, Cropp TA, Williams GJ. **Development of transcription factor-based designer macrolide biosensors for metabolic engineering and synthetic biology.** *ACS Synthetic Biology* 2018; **7(1):**227–39.

[19] Younger AKD, Dalvie NC, Rottinghaus AG, Leonard JN. **Engineering modular biosensors to confer metabolite-responsive regulation of transcription.** *ACS Synthetic Biology* 2017; **6(2):**311–25.

[20] Chen J, Sun S, Li CZ, Zhu YG, Rosen BP. **Biosensor for organoarsenical herbicides and growth promoters.** *Environmental Science and Technology* 2014; **48(2):**1141–7.

[21] Siedler S, Schendzielorz G, Binder S, Eggeling L, Bringer S, Bott M. **SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*.** *ACS Synthetic Biology* 2014; **3(1):**41–7.

[22] Siedler S, Stahlhut SG, Malla S, Maury JÔ, Neves AR. N**ovel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli.*** *Metabolic Engineering* 2014; **21:**2–8.

[23] Cayron J, Prudent E, Escoffier C, Gueguen E, Mandrand-Berthelot MA, Pignol D, Garcia D, Rodrigue A. **Pushing the limits of nickel detection to nanomolar range using a set of engineered bioluminescent *Escherichia coli.*** *Environmental Science and Pollution Research* 2017; **24(1):**4–14.

[24] Jha RK, Kern TL, Kim Y, Tesar C, Jedrzejczak R, Joachimiak A, Strauss CE. **A microbial sensor for organophosphate hydrolysis exploiting an engineered specificity switch in a transcription factor.** *Nucleic Acids Research* 2016; **44(17):**8490–500.

[25] Zhang C, Ye BC. **A single fluorescent protein-based sensor for *in vivo* 2-**

oxogluatarate detection in cell. *Biosensors and Bioelectronics* 2014; **54:**15–9.

[26] Kwon KK, Lee D-H, Kim SJ, Choi S-L, Rha E, Yeom S-J, Subhadra B, Lee J, Jeong KJ, Lee SG. **Evolution of enzymes with new specificity by high-throughput screening using DmpR-based genetic circuits and multiple flow cytometry rounds.** *Scientific Reports* 2018; **8(1):** 2659.

[27] Mahr R, von Boeselager RF, Wiechert J, Frunzke J. **Screening of an *Escherichia coli* promoter library for a phenylalanine biosensor.** *Applied Microbiology and Biotechnology* 2016; **100(15):**6739–53.

[28] Liu Y, Zhuang Y, Ding D, Xu Y, Sun J, Zhang D. **Biosensor-based evolution and elucidation of a biosynthetic pathway in *Escherichia coli.*** *ACS Synthetic Biology* 2017; **6(5):**837–48.

[29] Li H, Liang C, Chen W, Jin JM, Tang SY, Tao Y. **Monitoring *in vivo* metabolic flux with a designed whole-cell metabolite biosensor of shikimic acid.** *Biosensors and Bioelectronics* 2017; **98:**457–65.

[30] Frei CS, Wang Z, Qian S, Deutsch S, Sutter M, Cirino PC. **Analysis of amino acid substitutions in AraC variants that respond to triacetic acid lactone.** *Protein Science* 2016; **25(4):**804–14.

[31] Li Y, Qian S, Dunn R, Cirino PC. **Engineering *Escherichia coli* to increase triacetic acid lactone (TAL) production using an optimized TAL sensor-reporter system.** *Journal of Industrial Microbiology & Biotechnology* 2018; 789–93.

[32] Rosano GL, Ceccarelli EA. **Recombinant protein expression in *Escherichia coli:* Advances and challenges.** *Frontiers in Microbiology* 2014; **5:**1–17.

[33] Hecht A, Endy D, Salit M, Munson MS. **When wavelengths collide: bias in cell abundance measurements due to expressed fluorescent proteins***. ACS Synthetic Biology* 2016; **5(9):**1024–7.

[34] Raman S, Rogers JK, Taylor ND, Church GM. **Evolution-guided optimization of biosynthetic pathways.** *Proceedings of the National Academy of Sciences* 2014; **111(50):**201409523.

[35] Gao H, Kim IW, Choi JH, Khera E, Wen F, Lee JK. **Repeated production of l-xylulose by an immobilized whole-cell biocatalyst harboring l-arabinitol dehydrogenase coupled with an NAD+ regeneration system.** *Biochemical Engineering Journal* 2015; **96:**23–8.

[36] Zhang Z, Zhang R, Zou L, McClements DJ. **Protein encapsulation in alginate hydrogel beads: Effect of pH on microgel stability, protein retention and protein release**. *Food Hydrocolloids* 2016; **58:**308–15.

[37] Roberts TM, Rudolf F, Meyer A, Pellaux R, Whitehead E, Panke S, Held M. **Identification and Characterisation of a pH-stable GFP.** *Scientific Reports* 2016; **6:**1–9.

[38] Meyer A, Pellaux R, Potot S, Becker K, Hohmann H-P, Panke S, Held M. **Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors.** Nature Chemistry 2015; **7(8):**673–8.

[39] Song Y, Li J, Shin HD, Du G, Liu L, Chen J. **One-step biosynthesis of α-ketoisocaproate from l-leucine by an *Escherichia coli* whole-cell biocatalyst expressing an l-amino acid deaminase from *Proteus vulgaris.*** Scientific Reports 2015; **5:**1–11.

[40] Wei G, Ma W, Zhang A, Cao X, Shen J, Li Y, Chen K, Ouyang P. **Enhancing catalytic stability and cadaverine tolerance by whole-cell immobilization and the addition of cell protectant during cadaverine production.** *Applied Microbiology and Biotechnology. Applied* 2018; 1–11.

[41] Walser M, Pellaux R, Meyer A, Bechtold M, Vanderschuren H, Reinhardt R, Magyar J, Held M. **Novel method for high-throughput colony PCR screening in nanoliter-reactors.** *Nucleic Acids Research* 2009; **37(8):** e57.

[42] Barker DF, Campbell AM. **Use of bio-lac fusion strains to study regulation of biotin biosynthesis in Escherichia coli.** *Journal of bacteriology* 1980; **143(2):**789–800.

[43] Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. **Engineering and characterization of a superfolder green fluorescent protein.** *Nature Biotechnology* 2005; **24(1):**79.

[44] Paulsson J. **Summing up the noise in gene networks.** *Nature* 2004; **427(6973):**415.

[45] Anderson Jc, Dueber JE, Leguia M, Wu GC, Goler JA, Arkin AP, Keasling JD. **BglBricks: A flexible standard for biological part assembly.** *Journal of Biological Engineering* 2010; **4(1):**1.

[46] Salis HM. **The ribosome binding site calculator.** *Methods in Enzymology* 2011; **498:**19-42**.**

[47] Daâssi D, Rodríguez-Couto S, Nasri M, Mechichi T. **Biodegradation of textile dyes by immobilized laccase from *Coriolopsis gallica* into Ca-alginate beads**. *International Biodeterioration and Biodegradation* 2014; **90:**71–8.

[48] Smanski MJ, Bhatia S, Zhao D, Park Y, Woodruff BA, Giannoukos G, Ciulla D, Busby M, Calderon J…Gordon DB. **Functional optimization of gene clusters by combinatorial design and assembly.** *Nature Biotechnology* 2014; **32:**1241

CHAPTER THREE – INCREASING THE CELLULAR DEMAND FOR BIOTIN AS A SELECTION SYSTEM

For confidentiality reasons, aspects of this chapter have been modified,

Abstract

Although auxotrophic selection is a conventional approach for strain engineering, synthetic auxotrophy is only recently emerging in the field of synthetic biology. We construct here two selection systems for biotin biosynthesis by genetically modifying the host strain so that the resulting phenotype itself serves as selection pressure. In both cases, this selection pressure is due to increased cellular demand for the essential target metabolite or its immediate precursor. We use this method of synthetic auxotrophy to screen two different libraries and the effect of *E. coli*'s natural mutation rate.

INTRODUCTION

Modularity and bottlenecks

To generate high titers of a target molecule in metabolic engineering, the individual steps of a biosynthesis pathway often require fine-tuning, especially for non-native products or 1000-fold increases to *in vivo* concentrations. Modularity is a conducive approach to addressing such targets: expression constructs can be readily exchanged and tested for increase in product and pathway intermediates. This identifies and addresses engineering bottlenecks in a rapid, iterative manner. In synthetic expression constructs, several variables affect a pathway module. On the sequence level, transcription rates are determined by promoter strength, concentration of inducer (if relevant), copy number of the carrying plasmid, codon optimization, and the rate of translation initiation based on RBS sequence[8]. On a metabolic level, optimal transcription levels can nevertheless lead to poor production due to lack of precursor or cofactor availability as well as the genetic background of the cell factory. Generally it is desired to have a well-balanced metabolic pathway that achieves the optimum capacity of product formation. An imbalance is often characterised by low yields and can be due to low enzyme activity, poor expression, or low precursor availability (Figure 1).

Although low flux and pathway bottlenecks pose an engineering challenge, they can be used as a powerful selection tool. For example, low enzymatic activity can be compensated by high substrate availability (Figure 1B). Poorly expressed enzyme constructs can be used to identify pathway variants with high precursor concentration. Here, we construct a modular pathway for biotin biosynthesis (Figure 1A) and use a bottleneck BioB module to select variants that supply high amounts of precursor DTB.

[8] Additional variables, such as mRNA stability, folding efficiency, chaperone availability, or enzyme degradation, are not taken into account here as they are usually not considered part of the “toolbox” of metabolic engineering. This suitability of this approach is discussed elsewhere (22).

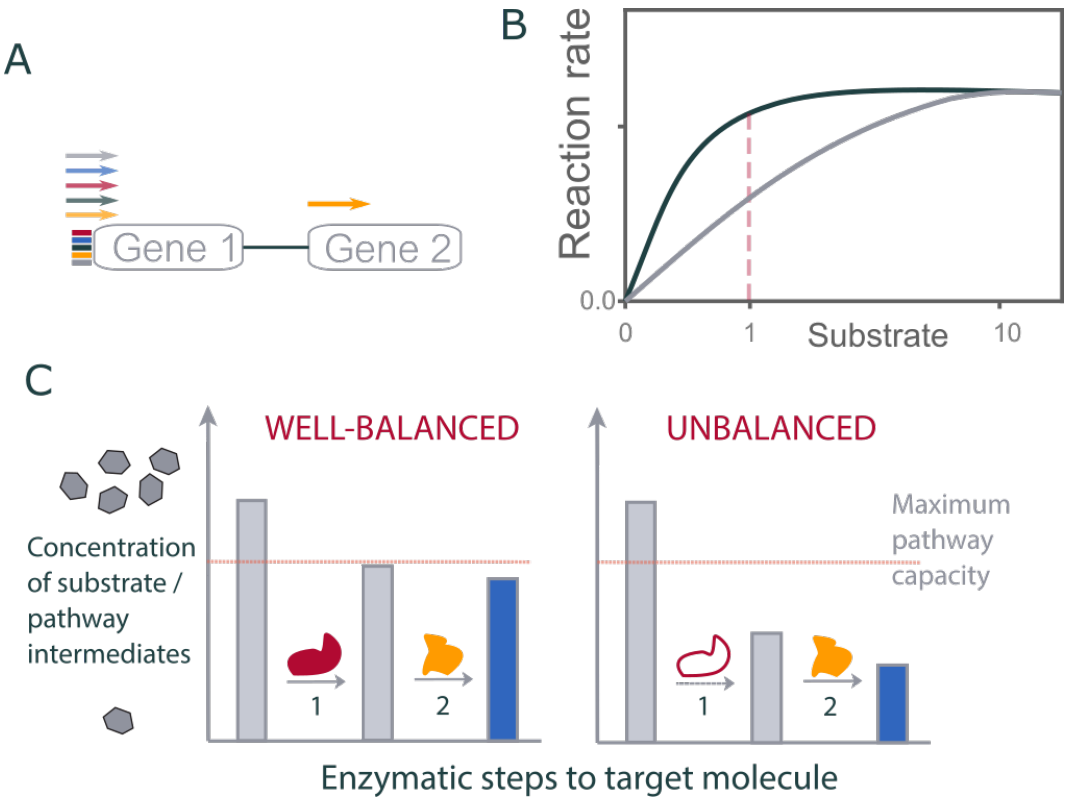


Figure 1: Factors contributing to pathway balance

A: Example of the biotin biosynthesis pathway as a modular, synthetic expression construct. Two steps of the pathway to DTB are shown. To identify well-producing variants, an RBS and promoter library with expected diversity 272 explores the expression landscape of Gene 1, while Gene 2 is constitutively expressed.

B: Idealised Michael-Menten curves explain the relationship between substrate availability and reaction rate. The darker line signifies an optimally expressed pathway step with a high rate of turnover. The light grey signifies an enzyme with suboptimal expression and therefore a poor reaction rate. At low substrate concentration (red dotted line), ~50% of V_{max} is achieved. This low catalytic rate can be compensated by high substrate availability.

C: A schematic of a simplified, two-step pathway is shown on the left under balanced conditions, where initial substrate (grey bar) availability for enzyme 1 (red, filled in) is higher than the maximum conversion rate as indicated by dotted line. Overall a high flux occurs through the pathway with only some loss of metabolite at enzyme 2 (yellow) with a high concentration of final product (blue bar). In the unbalanced pathway (right), enzyme 1 (red outline) is badly expressed or inherently poorly active so that only partial conversion of the substrate occurs. Although enzyme 2 has a higher capacity for conversion, the low substrate availability means that the overall yield of the pathway is low (blue bar).

Expanding *E. coli's* biotin auxotrophy

Strains with auxotrophy for the target molecule can be desirable in strain engineering. In a fermentation setting, auxotrophic dependency can replace antibiotics to maintain a production plasmid. This may be preferable as large-scale use of antibiotics is expensive and prone to mutation and cultivation of persister cells (1). Further, use of antibiotics can hinder regulatory approval for some processes and may also spread resistance genes into the environment (2). Finally, auxotrophic plasmid maintenance systems can provide genetic engineering tools for "safe for consumption" microorganisms such as the recently published *ΔthyA* strain of lactic acid bacteria (3). Synthetic auxotrophies[9] are underutilized in metabolic engineering despite promising results for targets such as 1-butanol (4, 5). Here we showcase the utility of synthetic auxotrophy as a selection system for metabolic engineering.

Biotin is essential for central carbon metabolism as it acts as a co-factor for the formation of malonyl-CoA, mediated by the bifunctional BirA ligase (cf. Introduction. In an auxotrophic strain (i.e. one incapable of synthesising the biotin moiety), biotin must be supplemented for growth. However the cellular demand for biotin is extremely low (6), much below production levels required industrially. Previously (7–10), mutations of the *E. coli* BirA sequence were generated and characterised. Of particular interest were those mutations which increase the intracellular demand for biotin and generate more auxotrophic strains. As BirA mediates biotin distribution in the cell, decreased binding efficiency or reduced catalytic ligase ability would lead to less optimal biotinylation of the AccB subunit and subsequently, a growth reduction (Figure 2 (11–14)). The growth reduction resulting from a poorly functioning BirA can be rescued by higher levels of intracellular biotin. We utilise a highly auxotrophic BirA mutation I207S to address the bottleneck BioB enzyme. We identify a plasmid with increased expression stability and interrogate a metagenomic library for heterologous sequences.

[9] A *synthetic auxotrophy* is a molecule dependency generated by rational strain engineering

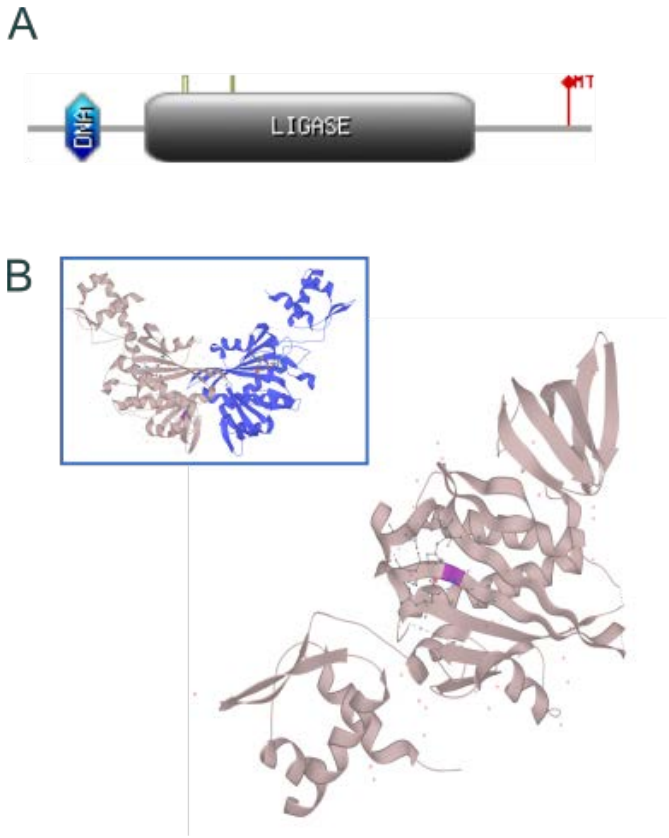


Figure 2: The bi-functional BirA protein

A - Functional domains of the *E. coli* BirA sequence, visualised at PROSITE (11). Shown is the ligase domain that transfers biotin to proteins, the DNA binding site, two biotin binding sites, and the mutation in red.

B - Crystal structures (12) were obtained from uniprot with mutation highlighted in pink (visualisation via litemol (13). Inset shows crystal structure of dimerised complex after biotin binding (14).

MATERIALS AND METHODS[10]

Bacterial strains and plasmids

Four plasmid-based metagenomic expression libraries were used in this study (Ab01, F01, F03, Granja Pig) which were generated from soil or fecal samples. Construction of the libraries as detailed (15) included the isolation of environmental DNA from the respective samples, sonication to generate 1-3 kB fragments from obtained DNA, and blunt-ended cloning into expression plasmid pBS18.

Table 2: Plasmids constructed in Chapter 2

Plasmid	Construction
pBS18	pZE21 (15)
pBS412	Amplification of kanamycin resistance cassette from pGEN49 (15) (oBS1283, oBS1284) Amplification of *bioB* from E. coli MG1655 genomic DNA Amplication of LacI and pSC101 region from pZS4Int-LacI (EXPRESSYS) (oBS1279, oBS1280)
pBS479	Addition of ampicillin resistance cassette (oBS886, oBS887)
pBS679	Addition of HindIII site with oBS1587

Chemicals

The HABA reagent (4'-hydroxyazobenzene-2-carboxylic acid) was obtained from Sigma-Aldrich. Streptavidin was obtained from Chemical Point.

Optimised MAGE protocols

The following protocol is modified from the published version (16).
In rich 2xYT media with kanamycin, an overnight culture of the desired strain harbouring MAGE plasmid pBS136 is used to inoculate a fresh media culture with kanamycinat a 1:100 dilution. At cuvette OD600 = 0.5 - 0.6, protein expression is induced with 2% arabinose.
After 30 min, culture is cooled, washed twice in ice-cold MQ, and 45 µl of cells used for electroporation with 10 µM oligo in a 1 mm cuvette at 1800 V. After recovery in SOC media, additional rounds of MAGE are carried out as desired.

[10] For general methods, please see Chapter 2.

BirA MT strains were generated with multiple rounds of MAGE using mOBS105 and confirmed with sequencing.

HABA assay for biotin equivalents

To make the HABA mix, 2 mg/mL streptavidin is dissolved in 7.5 mL 1x PBS and mixed with 1.5 mg/mL HABA reagent dissolved in 2.5 mL 10 mM NaOH. This is mixed 1:1 with 100 µL cell supernatant from production cultures. Standards between 5-25 µM are included. Absorbance is measured at 500 nm.

Confirming or reverting BirA MT

The biotin requirement of a BirA MT strain is higher than biotin concentrations generally available in LB media. To check against loss of the mutation, strains are streaked out on LB agar plates and LB agar with 0.1 mM biotin. Colony formation should only be observed on the latter. To revert the BirA MT to the WT state, a single round of MAGE was performed with moBS314. Transformants were plated onto LB agar as a selection for successful mutation reversion.

RESULTS

Poorly expressed pathway step as a selection

Complementing biotin synthesis with a BioB plasmid

To control biotin biosynthesis in *E. coli*, we knocked out the Δbio operon and replaced biotin synthase functionality with one of two BioB-expressing plasmid, pBS412 or pBS679. The differences between these two plasmids are i) pBS679 is ampicillin-resistance whereas pBS412 is kanamycin-resistant for compatibility with secondary plasmids, and ii) pBS679 contains a HindIII restriction site to increase clonability. Plasmid expression removes BioB transcription from feedback repression. Feeding the immediate precursor DTB allowed us to control turn-over of the enzymatic step. To allow fine-tuning of the system, BioB expression was controlled by an IPTG-inducible T5lacO promoter.

Both plasmids were characterised in strain BS1575 (Figure 3). In the experiments shown, production by both plasmids is comparable (Figure 3B-C, 0.9±0.01 compared to 1.3±0.02 mg/L). However, meta-analyis of 283 experiments with pBS679 and 98 experiments with pBS412 show an improvement for plasmid pBS679[11]. To streamline experimental parameters, only production tests at 0.01 mM IPTG and 100 mg/L DTB were included in the analysis, reducing the experimental number to 37 for pBS679 and 13 for pBS412; the remaining variable is production strain[12]. The mean titer for pBS679 resulting from 105 biological replicates is 1.25±0.6 mg/L (range [0.27; 2.69]); mean titer for pBS412 from 38 biological replicates is 0.86±0.5 mg/L (range [0.12; 1.54]). OD normalisation did not significantly affect the comparison. It was concluded that pBS679 not only gives higher titers but also produces in a more stable manner (CV = 48% versus 63%), although it is unclear why this is the case.

[11] From internal database.

[12] Assuming no differences due to experiment, equipment, etc.

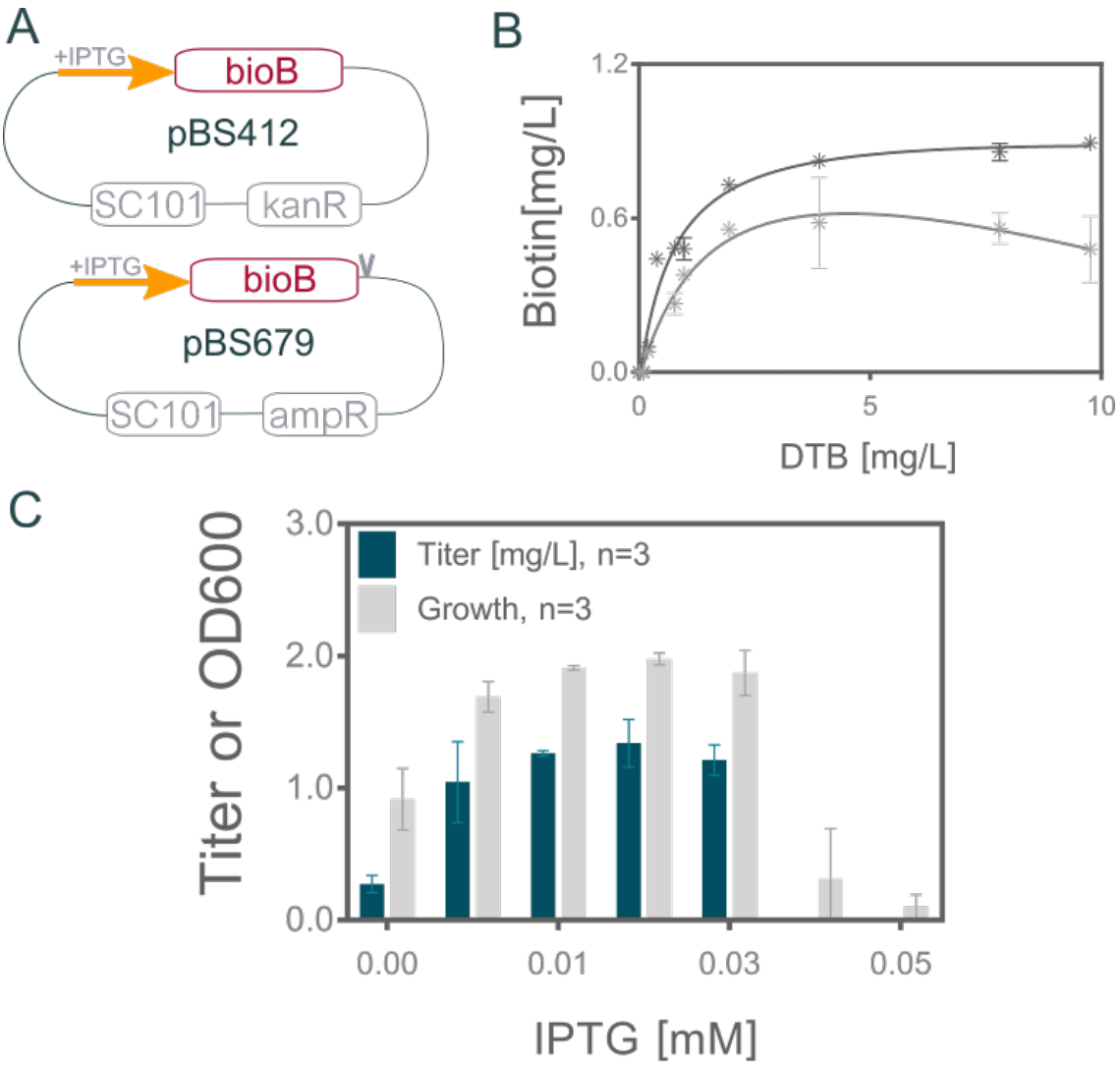


Figure 3: Plasmid architecture and activity of two BioB plasmids

A - Simplified schemes comparing the architecture of pBS412 and pBS67, both IPTG-inducible expression plasmids for BioB. The main difference lies in the antibiotics resistance marker and the addition of a restriction site following BioB in pBS679.

B - Strain BS1575 expressing BioB plasmid pBS412 was characterised for conversion of substrate DTB to biotin under set induction conditions (0.01 mM IPTG). Shown is biotin titers of a small-scale production experiment of a WT (light grey) and a mutant strain with improved FeS cluster supply (darker grey). The mutation both increases and stabilises the titer. Even at the lowest DTB concentration, the mutation increases the biotin titer. We exploit this for selection later in the chapter.

C - Optimal IPTG induction was determined for BioB plasmid pBS679 in BS1575. At excess DTB concentrations, the effect of IPTG concentrations between 0 and 0.05 mM were tested, indicating an optimal induction at 0.01 to 0.03 mM. High expression levels of BioB are toxic to the cell.

A second aspect of these BioB plasmids is toxicity at high expression, observable by reduced end-OD and cell viability. This can be seen at concentrations of IPTG > 0.3 mM in Figure 3C. A previously obtained expression mutant enables both higher catalysis and higher expression / lower toxicity (Figure 3B: WT K_D = 0.7669 with R^2 = 0.97; MT K_D = 1.884 R^2 = 0.986)[13]. This expression mutant was found to have improved iron-sulfur cluster supply, indicating that the bottleneck of BioB expression and the toxicity at high expression is involved with supplying or regeneration of iron-sulfur clusters.

BioB modularity for selection

A heterologous BioB from *Pseudomonas putida* (strain ATCC 47054 / DSM 6125 / NCIMB 11950 / KT2440) was inserted into the pBS679 backbone to generate pBS511 and characterised for activity in strain BS1575. To determine effect of expression and substrate on growth behaviour, strains were provided with high (10 mg/L = 47 µM), low (1 mg/L = 4.7 µM), or endogenous (assume <1 nM) DTB. Each condition was tested at 8 induction conditions [10, 1, 0.1 … 0]. No growth was observed for conditions less than 1 mg/L DTB or less than 0.1 mM IPTG, which is in stark contrast to plasmid pBS412 and pBS679. For these two plasmids in Δ*bioB* strains with the native pathway up to DTB intact, endogenous DTB levels and leaky expression (0 mM IPTG) is sufficient for full growth (data not shown). Activity of plasmid pBS511 was also tested on plates, where strains required ~2.5 mg/L DTB (10 µM) for full growth (data not shown).

[13] Fit and K_D calculated in Prism using one-site binding model. Non-specific binding component was disregarded, as the limit of detection of biotin bioassay is too high to determine background biotin levels.

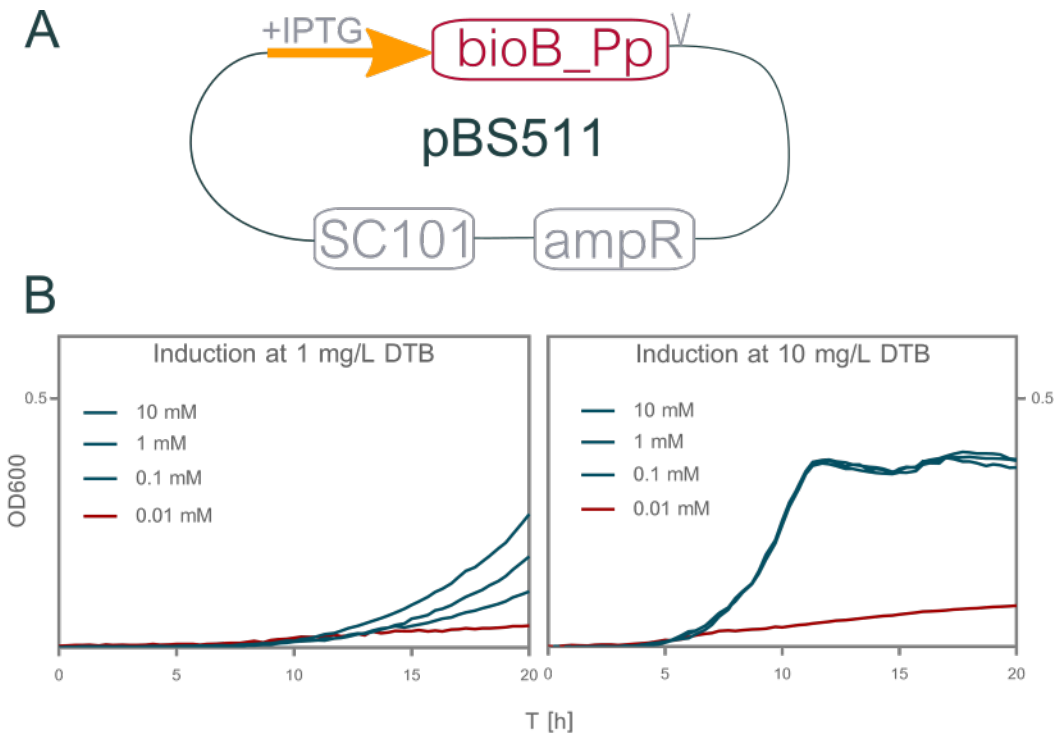


Figure 4: Controlling complementation using plasmid pBS511

A - A simplified graphic of the pBS511 shows the pBS679 backbone (pSC101 origin of replication, *ampR* ampicillin resistance, IPTG-inducible Lac-5 promoter, HindIII site after gene insertion site) expressing heterologous BioB from *P. putida.*

B - Plasmid pBS511 can complement a BioB KO strain to full growth only under conditions of high DTB availability (10 mg/L) and high induction (0.1 mM and above). Shown is growth (OD600) monitored over 24 hours.

Comparing the protein sequences for *E. coli* and *P. putida* BioB indicated a 67% similarity with no differences in iron-sulfur cluster binding residues[14]. The causative difference in complementation may be ascribed to RBS strength, which is reduced 19-fold in the *P. putida* sequence (33,000 AU vs. 1764 AU). This indicates that the high DTB demand of pBS511 is due to low translation of the BioB sequence (Figure 1, above).

The results shown in Figure 4 set up conditions to use growth-based selection of high DTB levels. We exploit this to identify high DTB pathways from a plasmid library of DTB producers. We generated a modular pathway plasmid with constitutive expression of *bioFAD* and a second ORF with heterologous *bioC* -

[14] Clustal Omega (23)

bioH from an unnamed organism (undisclosed), controlled by a library of promoter and RBS strengths (expected library size 272). As pBS511-complementation for BS1575 only occurs at high DTB availability, we hoped to identify the best promoter-RBS combination. The selection was carried out by electroporating the plasmid library with either pBS679 or pBS511 into BS1575 and plating out on minimal MOPS. Based on colony formation on LB (not shown) and selection plates (Figure 5A), the library with pBS679 had a total library size of 15-25 K. On selective conditions (MOPS with ampicillin and kanamycin, 0.01 mM IPTG), 10-12 K complemented auxotrophy. This indicates that between 40 to 80% of the plasmid library can complement the DTB requirement of pBS679 (< 1nM). The library with pBS511 had a total library size of 10-12 K. On selective conditions, (MOPS with ampicillin and kanamycin, 0.1 mM IPTG), 250 members fulfilled the much higher DTB demand of the plasmid (2.5 mg/L at 0.1 mM IPTG on plates). This represents 2 - 2.5% of library.

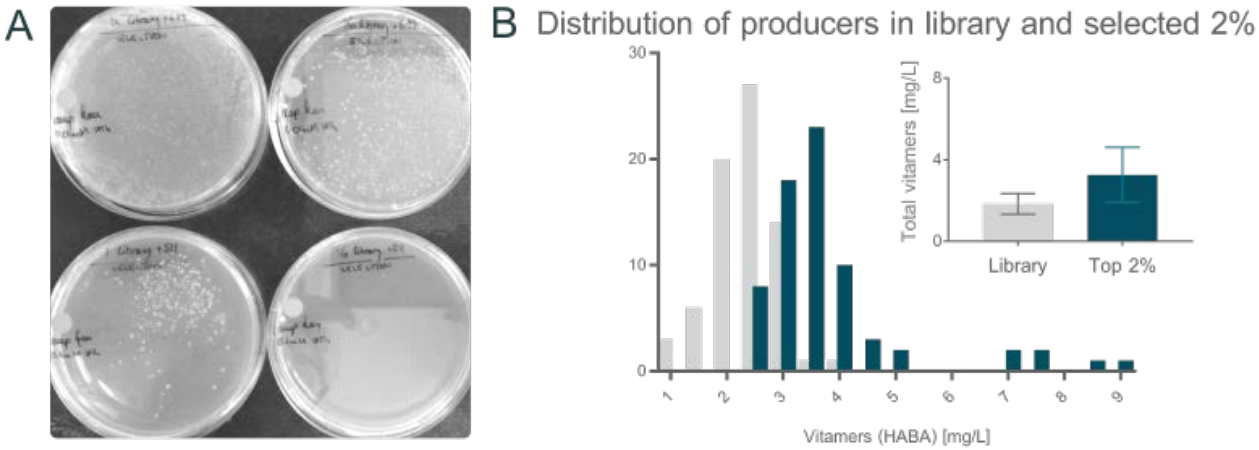


Figure 5: Selection and characterisation of a DTB library

A - Strains were transformed with ligation mix of library plasmids and good or poor BioB. Shown is colony formation at 1-fold and 10-fold dilution of the washed recovery mixture for library containing pBS679 (top row) or pBS511 (bottom row) on selection plates (minimal MOPS). The poor BioB valve strains only form colonies at the highest concentration of cells, approximately 100-fold less than the normal strains.

B - Characterisation of the whole library compared to those that enabled growth with pBS511 indicate a shift of the distribution towards higher production ($p<0.001$, unpaired t-test with Welch's correction), even though the average production of each set is within error. Note that the whole library is 50-fold larger.

To confirm that the selected 2% of library members were indeed the top performers in terms of DTB production, 72 colonies of each transformation from selective plates were tested for concentrations of biotin equivalents[15] in the supernatant. Based on the results (Figure 5B), the set of selected library members contained the highest producers of biotin equivalents and its highest 10% percentile is higher than any tested members of the whole library. Based on the production data, growth of a strain with pBS511 is possible with 2.5 mg/L extracellular biotin equivalents or higher[16]. Theoretically, this is equivalent to 23% of the total library (16 out of 72 colonies produces more than 2.5 mg/L biotin equivalents). This high-producing segment of the library was enriched more than 6-fold in the selection.

[15] DTB, biotin, oxidised biotin species.
[16] Which correlates with the DTB requirement on minimal plates.

Increasing the demand for biotin selects stable expression variants of BioB

From literature, we identified a mutant BirA sequence that increases the cellular demand for biotin. The mutation I207S was transferred into strain BS1575 using MAGE to generate BS1601. The cellular biotin demand was characterized. The increased biotin requirement was used to select for more stable expression plasmids for BioB.

Biotin requirement of BirA strains in liquid media and plates

The biotin demand of B1601 was characterised in liquid media (Figure 6) and on minimal agar plates (Figure 7A). Both conditions were supplemented with [0; 10 µM] biotin. After 4 days incubation on plates at 37 °C, no cells grew at 0 mM biotin whereas good growth was observed at 2.5 µM biotin and above on the plates. Some colony formation occurred at 1 µM indicating this concentration as the cut-off for auxotrophy complementation. The biotin demand in liquid for growth within 24h was much higher: full growth was observed only at 500 µM extracellular biotin and an auxotrophic cut-off above 7.5 µM. The shorter time interval was chosen due to limitations of the liquid experiment set-up: the kinetic reads are carried out on 200 µL growth culture, which can evaporate at elongated experimental conditions.

The auxotrophy of a BirA mt strain could also be complemented by a BioB expression plasmid in the presence of precursor DTB. The plating experiment was repeated with strain BS1601 carrying plasmid pBS679 (BS1874). Both parameters of the BioB reactions (precursor concentration and BioB induction level) were tested for auxotrophy complementation. At ideal induction conditions (0.01 mM IPTG), colony formation was characterised as a function of DTB availability (Figure 7B). Full complementation, *i.e.* identical growth on minimal agar plate as with supplemented biotin, was seen at 10 µM DTB. Concentrations of 2.5 µM DTB allowed some colony growth; 1 µM DTB allowed none.

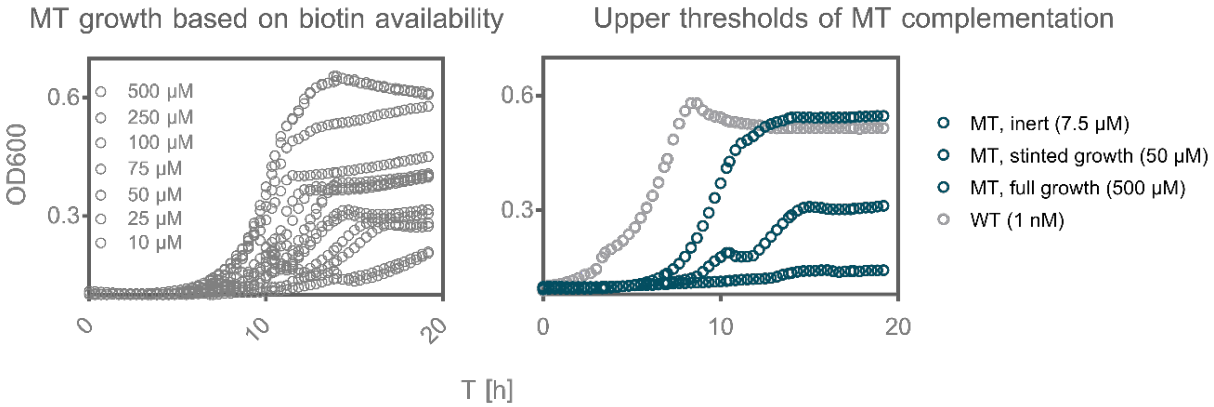


Figure 6: Biotin requirement of BirA I207S mutant strain

Left - Growth curves of strain BS1601 (BirA I207S, *Δbio*) in minimal media supplemented with biotin. Shown are two replicates for each concentration stated. Highest two growth curve corresponds to 500 µM, the 250 µM, etc.

Right - Three thresholds of growth are identified for strain BS1601: inert (no growth) up to 7.5 µM biotin, partial complementation (stinted growth) when supplemented between 10 and 50 µM biotin, and full complementation (full growth) at a minimum of 500 µM supplemented biotin. In grey is a WT strain (BS1575, BirA WT, *Δbio*) for comparison showing full complementation at 1 nM biotin. Each curve represents the mean of four replicates.

Alternatively, in conditions of excess precursor (Figure 7C), colony formation of a BirA mt strain could be characterised based on induction of BioB enzyme. Triplicate cultures of strain BS1874 were spotted (5 µL0at different dilutions at CFUs of 10^6/mL,10^5/ml, *etc.* on minimal MOPS plates with ampicillin and 50 µM DTB containing varying amounts of IPTG [0; 0.1 mM]. At no induction, the leakiness of the T7-lac promoter was sufficient to complement auxotrophy yet reduced growth is visible at lower cell concentrations (left spots compared to right). An increase in induction led to better growth in all spots up to 0.01 mM IPTG. At higher concentrations of IPTG, spot formation started to suffer due to expression toxicity; almost no survival was seen at 0.1 mM IPTG. This mirrored the behaviour of pBS679 in a BirA WT strain (see above).

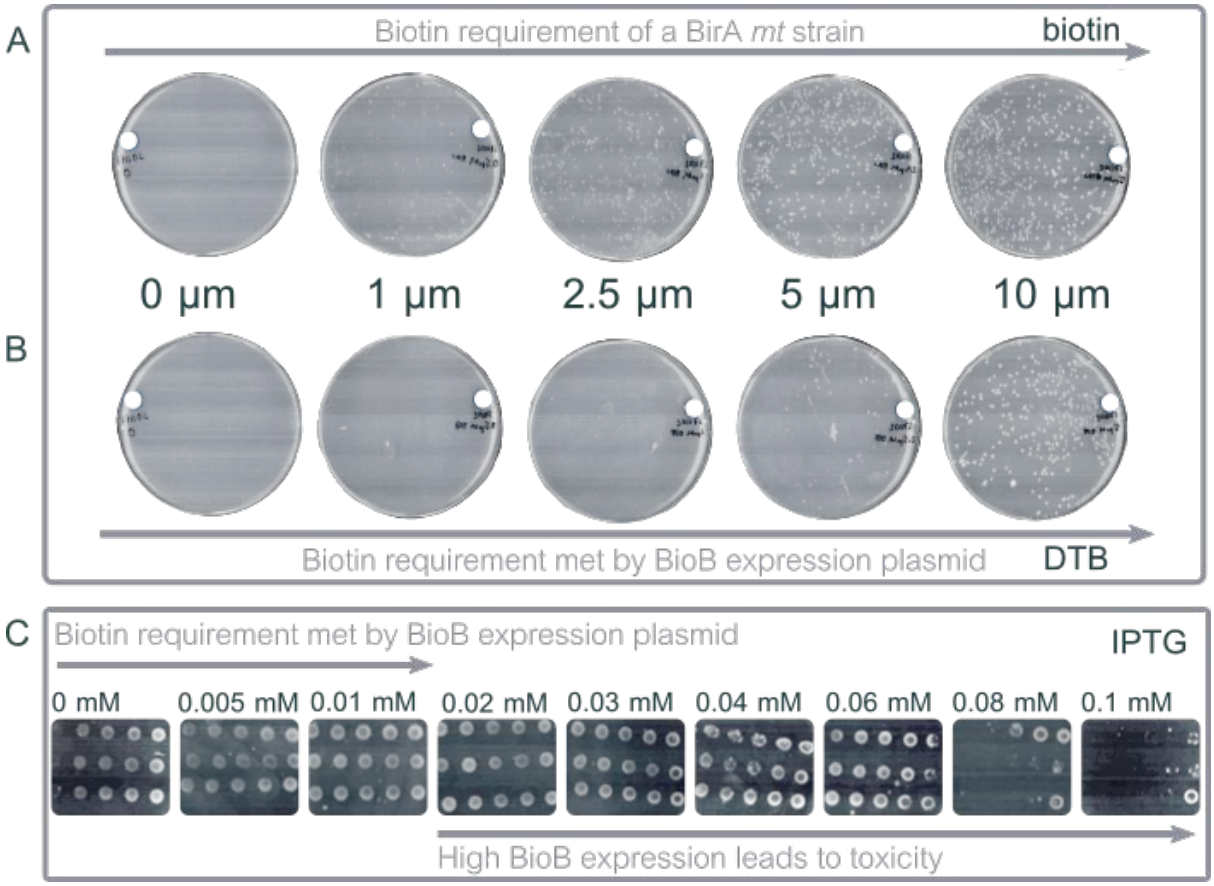


Figure 7: Controlling auxotrophy complementation using plasmid pBS679

A - The growth defect of a BirA mt strain can be alleviated by the addition of biotin at 1-2.5 µM, quantified by colony formation on minimal agar plates with supplemented biotin.

B - Further the strain can grow with supplemented biotin precursor DTB at 2.5 µM if expressing an inducible BioB plasmid at ideal induction of 0.01 mM IPTG.

C - BioB expression is toxic to the strain at high levels (above 0.02 mM IPTG, quantified by spot assay) when supplemented with 100 mg/L DTB.

The improved Fe-S cluster mutant shown above was previously isolated by plating BS1575 at high IPTG conditions to identify strains that could tolerate a higher BioB expression. We wanted to repeat this approach using a BirA strain as the previous selection round suffered from a high rate of false positives. Due to BirA auxotrophy, only functional BioB plasmids would allow for survival at non-toxic IPTG concentrations (0.01 mM) and additional expression mutants would allow for colony formation at 0.1 mM IPTG. Strain BS1874 was plated out on selective conditions (0.1 mM IPTG) at densities of 10^9 CFUs, leading to the formation of ca. 300 colonies (selection 1 in 3.3M, data not shown).

Out of 300 colonies, 200 were tested for biotin production in the presence of 100 mg/L DTB. 15% showed higher production than a non-selected strain (Figure 8A). These were counter-screened for intact BirA mutation by ensuring biotin requirement was still elevated. Colonies were tested for production and assessed for stability across replicates. Four hits were identified for confirmation, yet did not perform better than the control (BS1575 + pBS679) when cured and re-transformed with pBS679. Hence it was concluded that the increase in production seen in Figure 8C is due to mutations in the plasmid. Three plasmids were extracted and retransformed into BS1575. No increased titers were observed but a change in expression at different concentrations of IPTG was seen for plasmid pBS679* (Figure 8C). Sanger sequencing indicates an insertion resulting in a frameshift mutation for the last third of the gene sequence. However, as titers were not improved, we considered the hits from this selection round false positives.

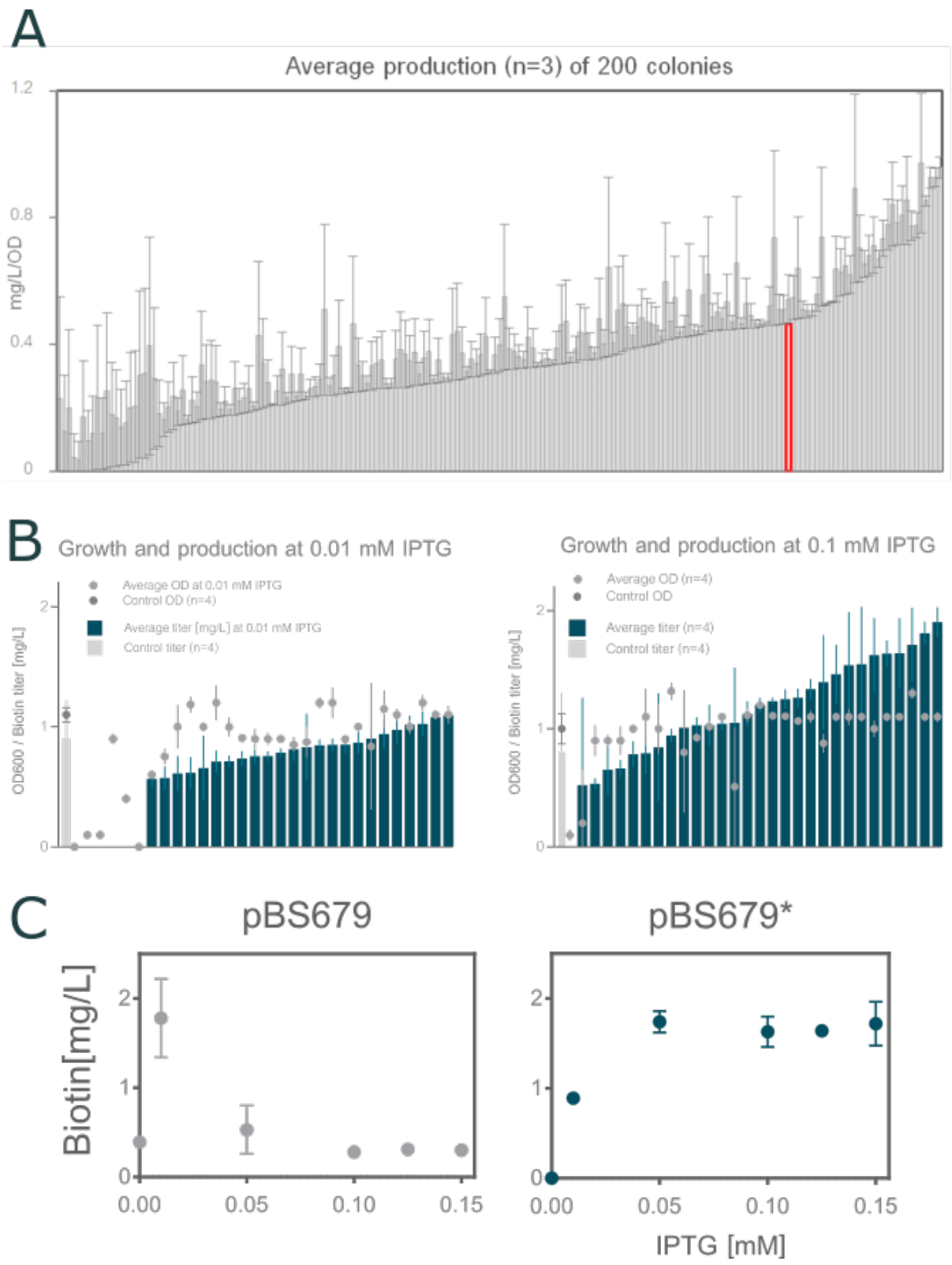

A
Average production (n=3) of 200 colonies
mg/L/OD
1.2
0.8
0.4
0
B
Growth and production at 0.01 mM IPTG
Average OD at 0.01 mM IPTG
Control OD (n=4)
Average titer [mg/L] at 0.01 mM IPTG
Control titer (n=4)
OD600 / Biotin titer [mg/L]
Growth and production at 0.1 mM IPTG
Average OD (n=4)
Control OD
Average titer (n=4)
Control titer (n=4)
C
pBS679
pBS679*
Biotin[mg/L]
0.00
0.05
0.10
0.15
IPTG [mM]

Figure 8: A biotin selection system using engineered auxotrophy

A - 200 colonies survived selection conditions (0.1 mM IPTG) and were tested for production. Biotin concentrations of supernatant were determined by bioassay using a Δ*bio E. coli* strain and sorted by lower significance cut-off (mean - stdev). Compared to control strain (red, a BirA mt not subject to selective plating), 30 colonies showed higher production.

B - Thirty colonies were chosen for further characterization based on increased titers with stable production. Shown is growth and biotin concentration for each strain at two induction conditions compared against control (grey bar, left) in each case. Note that strain phenotypes were reverted to BirA WT. Four strains were identified for further testing.

C – Three plasmids were successfully isolated from improved strains identified in B and freshly transformed into BS1575. They were tested for production but no improved titers were seen at 100 mg/L DTB and the tested IPTG gradients. However, a more robust plasmid pBS679* was identified with less variability between replicates.

Functional metagenomic selection

Functional metagenomic selection is the use of gene KOs to identify heterologous sequences to complement functionality. To interrogate a total of four 1 kB, plasmid-based metagenomic expression libraries for i) alternative BioBs sequences and ii) more efficient catalysis, we utilised a two-strain approach with BirA WT and MT strains.

In two rounds of transformation of the plasmid libraries into either the BirA WT or MT strain, a total of >800,000 individual transformants (CFU's on non-selection plates) were isolated. Transformation efficiencies were up to 7-fold higher for the BS1575 strain (700,000 CFUs versus 120 000 CFUs); we assume the slower growth of strain BS1601 impedes transformation rates. On selective plates (minimal MOPS with 1 µM DTB), colony formation indicated selection rates between 0.5 – 4% regardless of strain. We observed a high background rate on negative control plates (background strain with pBS679 plated onto minimal MOPS without DTB), which we ascribe to trace biotin contamination. Cross-talk of strains capable of biotin production lead to a halo-like formation around their colonies (Figure 9A). Colonies were restreaked and sequenced, identifying 16 biotin synthases from *Serratia spp.* and *Pseudomonas spp.* (Figure 9B). Further, 8 sequences showed the complementation phenotype but were not predicted biotin synthases following blastP and blastX analysis (17). No biotin synthases were isolated with the BirA MT strain. Although a total of 50 colonies appeared to complement functionality, sequenced revealed these as alternative BPL which removed the selection pressure of the screen (false positives).

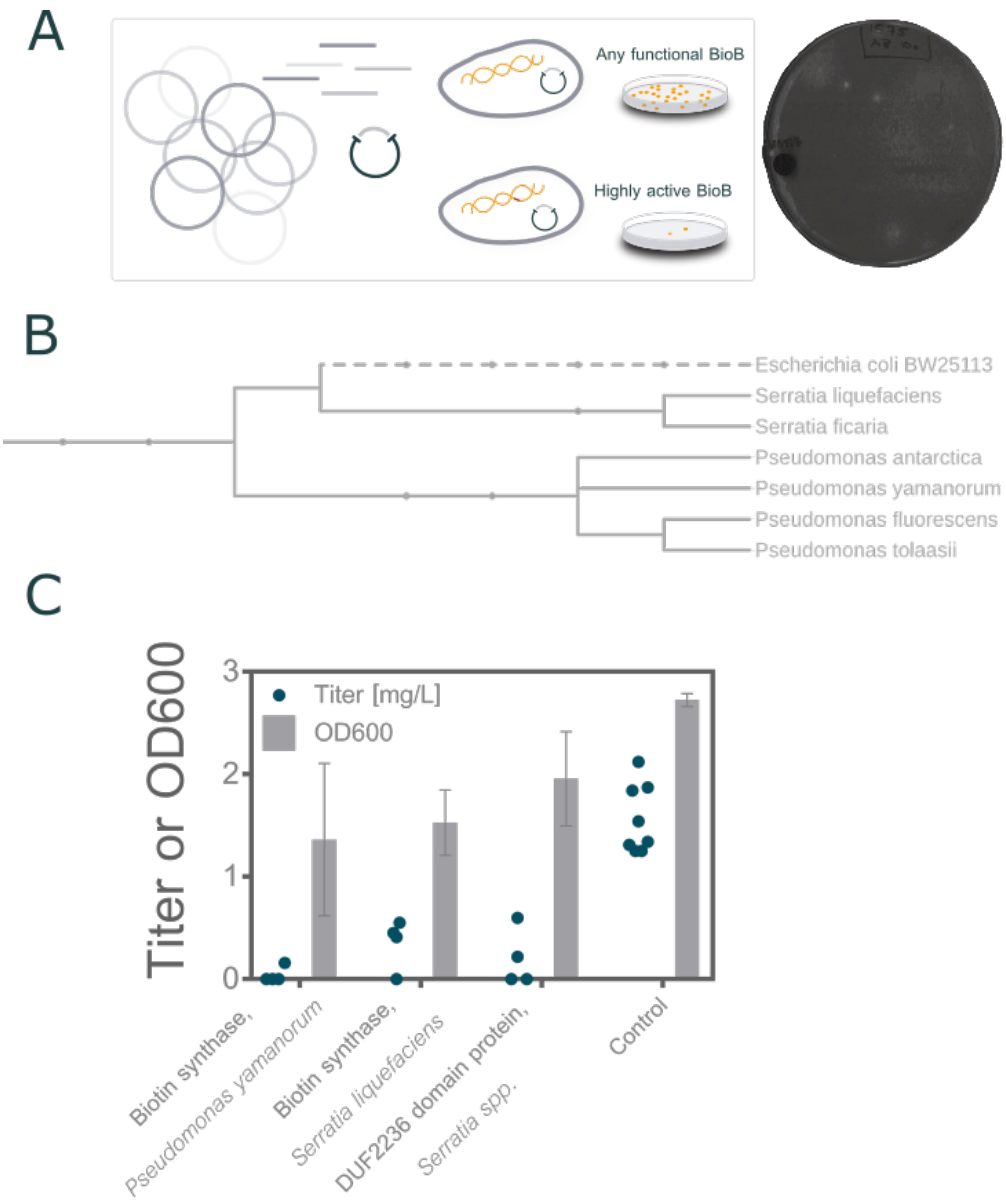


Figure 9: Functional metagenomic selection with BirA MT or WT strains

A – Overview of experimental set-up: Environmental DNA was shredded into 1 kB fragments and inserted into an expression backbone to generate the metagenomic libraries. These were transformed into Δ*bio* BirA WT or MT strain and plated at minimal DTB concentration to identify clones that could complement the respective strain auxotrophy. An unexpected result from the plating was cross-talk from high producers which allowed false positives to grow around the colony in a halo formation.

B – Phylogenetic tree of 16 BioB sequences identified in the BirA WT selection.

C – Phenotypic characterization of the sequences expressed in the metagenomic plasmid backbone was carried out in the presence of 100 mg/L to determine production potential of the plasmids in the BS1575 backbone. Three of the hits gave non-zero production values (n=4).

Three plasmids showed non-zero biotin titers when transformed into BS1575 and tested for production in the presence of 100 mg/L DTB (Figure 9C). To test production capacity in a suitable backbone with balanced RBS and optimal induction conditions, the two annotated BioB sequences were transferred into the pBS679 backbone. Cloning was successful only for the sequence from *Serratia liquefaciens* (pBS1406). The sequence did not outperform *E. coli* BioB (pBS679, Figure 10).

Using a functional metagenomics approach, we aimed to interrogate a library of metagenomic expression plasmids to identify i) *any* BioB sequences that could complement a Δ*bioB* knockout strain, and ii) improved BioB sequences that could complement a BirA WT under DTB conditions where *E. coli* WT could not. We identified 16 biotin synthases that functionally expressed in the *E. coli* host, but did not perform better than pBS679 when transferred to this plasmid backbone. Hence, none of these sequences would have been capable of complementing BirA auxotrophy under the selection conditions.

The selection scheme would benefit from two modifications. Firstly, a larger metagenomic expression library may contain a broader diversity of biotin synthases. Secondly, the low transformation efficiency of the BirA MT strain may impede larger metagenomic screens. Optimising of the transformation protocol for this strain may be advisable.

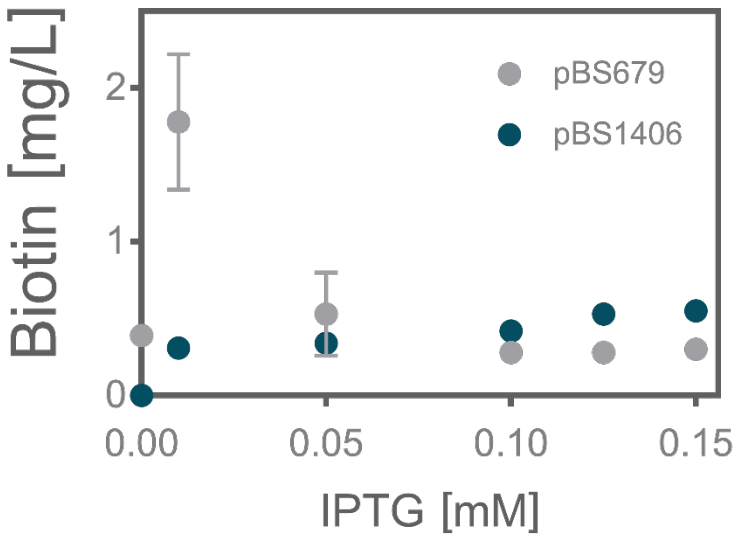


Figure 10: Production of pBS1406 compared to pBS679

Three biological replicates of strain BS1575 carrying either pBS679 or pBS1406 were tested for production based on IPTG induction in the presence of 100 mg/L DTB. Supernatant titers were analysed by bioassay. Error bars of strain carrying pBS1406 were too small to see.

DISCUSSION

Metabolic engineering aims to increase production and stability of cell factories. Understanding metabolite valves and modularizing pathway steps are common tools to reach this aim. Optimising flux through metabolite valves directs more resources from central carbon metabolism into the production pathway without negatively affecting growth and cellular upkeep (18–20). Modularity is used as a design concept to increase functionality, especially for the systematic biosynthesis of structurally related target compounds. For example, the engineering of glycosylation patterns for erythromycin production in *E. coli* (21) demands a methylation at C3 of each of the various substrates. In the work presented in this chapter, we utilise the tools of metabolic engineering (valves and modularity) yet reconfigure the main aim. We decrease strain production and through this, generate selection pressure.

The described selection systems were used to successfully interrogate different types of libraries. In the first part, a poorly expressed BioB produces biotin concentrations at the lower threshold of *Δbio* complementation. This construct was utilised to enrich a library of DTB producers and identify several high-producing variants. In the second part, a synthetically engineered auxotrophy finds plasmid mutations that stabilize production but do not increase titers. Further, the engineered auxotrophy is used to test a plasmid library of metagenomic fragments for improved BioB variants. No BioB sequences performed better than the native *E. coli* sequence and could only be isolated with a BirA WT strain with no auxotrophic selection pressure. The strain with engineered auxotrophy may have been too stringent for the library. Further, the BirA MT suffered from lower transformation efficiencies and a high rate of false positives.

It should be noted that engineered auxotrophy as a selection system may not be suitable for metagenomic selections. In our case, the mutant BirA selection pressure was easily evaded by the expression of heterologous BPL sequences. These false positives quickly overwhelmed the selection system. Nevertheless, we are confident that the functional landscape of the metagenomic libraries was explored, as a superior BioB would have complemented the increased auxotrophy of the BirA MT *and* generated the halo formation which was seen in the BirA WT screen.

BIBLIOGRAPHY

[1] Mülleder M, Campbell K, Matsarskaia O, Eckerstorfer F, Ralser M. ***Saccharomyces cerevisiae* single-copy plasmids for auxotrophy compensation, multiple marker selection, and for designing metabolically cooperating communities.** *F1000Research* 2016; **5(0):**2351.

[2] Vidal L, Pinsach J, Striedner G, Caminal G, Ferrer P. **Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*.** *Journal of Biotechnology* 2008; **134(1–2):**127–36.

[3] Li C, Dong H, Lu H, Gu X, Tian J, Xu W, Tian H. **Development of an antibiotic-free plasmid selection system based on thymine auxotrophy in *Lactococcus lactis*.** *Annals of Microbiology* 2015; **65(2):**1049–55.

[4] Shen CR, Liao JC. **Synergy as design principle for metabolic engineering of 1-propanol production in *Escherichia coli.*** *Metabolic Engineering* 2013; **17(1):**12–22.

[5] Wen RC, Shen CR. **Self-regulated 1-butanol production in *Escherichia coli* based on the endogenous fermentative control.** *Biotechnology for Biofuels* 2016; **9(1):**1–15.

[6] Choi-Rhee E, Cronan JE. **Biotin synthase is catalytic *in vivo*, but catalysis engenders destruction of the protein.** *Chemistry & Biology* 2005; **12(4):**461–8.

[7] Campbell A, Del Campillo-Campbell A, Chang R. **A mutant of *Escherichia coli* that requires high concentrations of biotin.** *Proceedings of the National Academy of Sciences* 1972; **69(3):**676–80.

[8] Campbell A, Chang R, Barker D, Ketner G**. Biotin regulatory (bir) mutations**

of *Escherichia coli.* *Journal of Bacteriology* 1980; **142(3):**1025–8.

[9] Barker DF, Campbell AM. **Use of bio-lac fusion strains to study regulation of biotin biosynthesis in *Escherichia coli.*** *Journal of Bacteriology* 1980; **143(2):**789–800.

[10] Chakravartty V, Cronan JE. **The wing of a winged helix-turn-helix transcription factor organizes the active site of BirA, a bifunctional repressor/ligase.** *The Journal of Biological Chemistry* 2013; **288(50):**36029–39.

[11] Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I**. New and continuing developments at PROSITE.** *Nucleic Acids Research* 2013; **41(D1):**D344–7.

[12] Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, Matthews BW. ***Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains.** *Proceedings of the National Academy of Sciences* 1992; **89(19):**9257–61.

[13] Sehnal D, Deshpande M, Vařeková RS, Mir S, Berka K, Midlik A, Pravda L, Velankar D, Koča J. **LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data.** *Nature Methods* 2017; **14:**1121.

[14] Smanski MJ, Bhatia S, Zhao D, Park Y, Woodruff BA, Giannoukos G, Ciulla D, Busby M, Calderon J…Gordon DB. **Functional optimization of gene clusters by combinatorial design and assembly.** *Nature Biotechnology* 2014; **32:**1241

[15] Genee HJ. **Towards Evolution-Guided Microbial Engineering - Tools Development and Applications (Ph.D. Thesis).** The Technical University of Denmark; 2015.

[16] Lennen RM, Nilsson Wallin AI, Pedersen M, Bonde M, Luo H, Herrgård MJ, Sommer MOAS. **Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with**

reduced off-target effects. *Nucleic Acids Research.* 2015; **(11):**e36.

[17] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. **BLAST+: architecture and applications**. *BMC Bioinformatics* 2009; **10(1):**421.

[18] Brockman IM, Prather KLJ. **Dynamic knockdown of *E. coli* central metabolism for redirecting fluxes of primary metabolites.** 2016; **48(0):**923–30.

[19] Farmer WR, Liao JC. **Improving lycopene production in *Escherichia coli* by engineering metabolic control.** *Nature Biotechnology* 2000; **18(5):**533–7.

[20] Solomon K V., Sanders TM, Prather KLJ. **A dynamic metabolite valve for the control of central carbon metabolism.** *Metabolic Engineering* 2012; **14(6):**661–71.

[21] Zhang G, Li Y, Fang L, Pfeifer BA. **Tailoring pathway modularity in the biosynthesis of erythromycin analogs heterologously engineered in *E. coli*.** *Science Advances* 2015; **1(4)**:1500077.

[22] Zhang J, Jensen MK, Keasling JD. **Development of biosensors and their application in metabolic engineering**. *Current Opinion in Chemical Biology* 2015; **28:**1–8.

[23] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M … Thompson JD. **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Molecular Systems Biology* 2011; **7(1):** 539.

CHAPTER FOUR – A COUNTER-SCREEN FOR TOXICITY SELECTIONS

This chapter is printed in its original form.

Abstract

In our biotin production strains, the bottleneck enzyme BioB is expressed under an IPTG-inducible promoter to control the trade-off between synthesis and toxicity. Over time, some of the strains have become IPTG inert in that they do not show the expected expression level at given induction conditions. We call this the “tuning phenotype” and have observed it caused by both plasmid- and genome-based mutations. To counter-screen for tuned strains, we generate a *bioB-gfp* transcriptional fusion and test for robustness in small-scale and shakeflask. Finally, we carry out a round of selection with the fusion plasmid and reduce the likelihood of false positives 3-fold.

INTRODUCTION

Selection systems rely on applying survival pressures to identify high producers. This can lead to false positives as mutated strains with loss-of-selection outgrow other strains. Loss-of-selection is attributable to the accumulation of mutations in production or selection modules. This is of particular relevance if PCR amplification is used, as the error rate of most polymerases is up to 10-fold higher than natural mutation (1), or if toxic genes are maintained on the plasmid. The growth advantage of losing a production or selection plasmid was quantified by an increase in growth rate by 30% in two independent studies (2,3). Without any selection pressure, plasmids were lost within 48h of continuous culture (2).

To address the issue of loss-of-selection and false positives, counter-selection strategies are often employed. One type of counter-selection is the addition of selectable markers as kill switches: *rpsL, galK, thyA, tetA*, *tolC*, or *sacB* have been utilised. After a selection hit has been identified, the intactness of the selection system can be confirmed by the effect of the lethal gene. Other options include colourimetric assays such as *xylE* expression (4), or the nucleoside

kinase activity of herpes simplex virus thymidine kinase (hsvTK) which can restore cell growth when provided with substrate dP (1). A second type of counter-selection is redundancy to safeguard against escapees, as already discussed in Chapter 1. A further example is where low product concentrations led to cell death. For example, a system was built to find novel proteases were the desired enzymatic activity was coupled to b-lactamase secretion and resulting antibiotic resistance (5).

In our set-up, selective pressure was generated by the toxicity of high BioB expression.
After identification of an expression mutant which reduced toxicity and increased titers (detailed in Chapter 3), further rounds of “toxicity selections” were carried out. However, the scope of loss-of-selection and resulting 100% false positive rate impeded any additional rounds of this set-up, as all colonies that survived the toxicity challenged showed the tuning phenotype. Instead of constructing a counter-selection, we attempted to build a counter-screen that merely monitored BioB activity. For this we generated a transcriptional fusion of *bioB* with the *gfp* gene. Fusions with fluorescent proteins are excellent monitors of transcriptional activity if their expression does not lead to cellular burden. They have been used extensively to study bacteria, including promoter activity (6) or heterogeneity of single cell expression (7). Here we present their use as a monitor for the tuning phenotype.

MATERIALS AND METHODS[17]

Bacterial strains and plasmids

Table 1: Plasmids constructed in Chapter 4

Plasmid	Construction
pBS683	Amplification of backbone from pBS679 Amplification of sfGFP from pBS046 (oBS1590, oBS1591)
pBS1058	Amplification of backbone from pBS679 (oBS2236, oBS2237) Amplification of sfGFP from pBS452 (obS2374, oBS2375)

Shake-flask cultivations

An overnight culture of relevant strains was started from single colony (biological duplicates) in 0.4 mL MOPS ampicillin with 1 µM biotin or DTB as appropriate. Using baffled flasks (Schott), 50 mL media (MOPS with ampicillin and 100 mg/L DTB) was inoculated from the overnight culture at a dilution of 1:50 and incubated at 275 RPM, 37 °C. Shake-flasks contained 0.01 mM IPTG for BioB plasmid cultures and GFP plasmid cultures, and 0 or 10 mM IPTG for BioB-GFP plasmid cultures. GFP plasmid cultures also contained 1 µM biotin. Samples were taken under sterile conditions at time-points 2, 4, 5.5, 7, 8.5, 10, and 24h and measurements of OD, fluorescence, and biotin concentrations in the supernatant were taken.

[17] For general methods, please see Chapter 2.

RESULTS

Tuning plasmids - a problem

Plasmids pBS412 and pBS679 expresses *bioB,* the bottleneck of biotin biosynthesis in *E. coli* (described in Chapter 3). The gene is expressed by a T5-lac promoter. Increasingly high IPTG induction of the construct led first to growth reduction and then non-viable cells. In a separate work, toxicity of BioB was used as a selection system where high cell amounts were plated out on toxic IPTG conditions. It was hypothesised that colonies that could grow under the condition either i) reduced toxicity of BioB or ii) increased expression capacity of the strain. This approach led to the identification of a "expression mutation" that increased expression and allowed for higher production from the plasmids via improved iron-sulfur cluster supply. BS1575 contains this mutation.

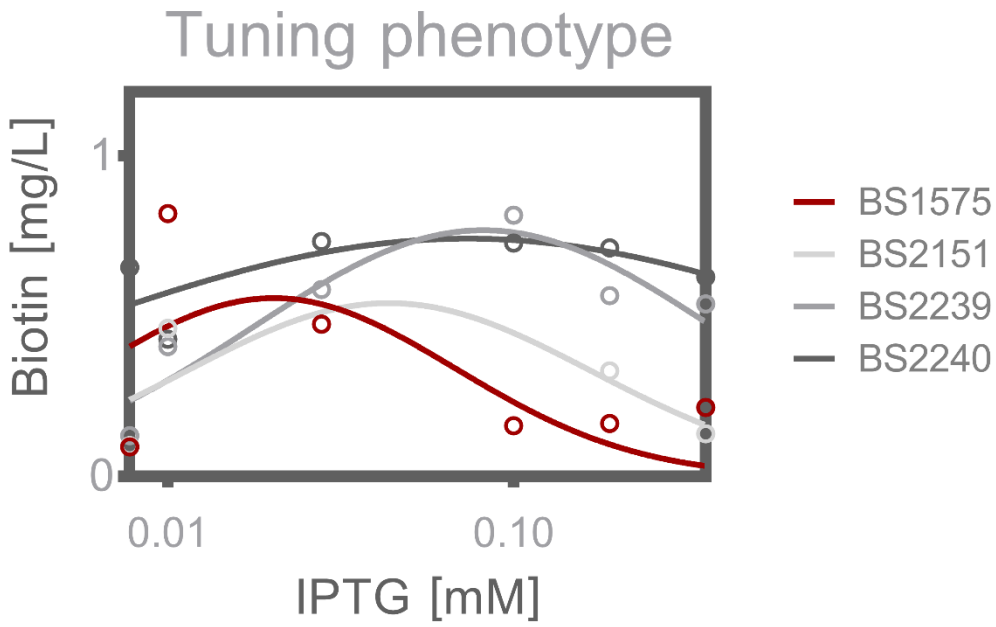


Figure 1: Optimal IPTG induction can shift in strains after a round of toxicity selection

Three daughter strains of BS1575 with pBS679 were obtained after a round of toxicity selection. Although none of them have higher titers than the control (in red), each strain is more resistant to IPTG induction. The concentration for optimal BioB expression (determined by titer) has shifted from 0.01 mM to 0.05 (light grey line) or 0.1 mM (darker grey lines) dependent on strain.

However, the selection led to a high rate of false positives wherein strains could grow on high induction conditions but did not have a higher production capacity (Figure 1). These hits have tuned expression of the toxic sequence so that they are less responsive to IPTG. Upon further characterization, 100% of "hit" plasmids show the tuning phenotype, which made further rounds of toxicity selections challenging.

To explore the relationship between IPTG and the tuning phenotype, we exposed strain BS1574[18] with pBS679 to increased induction conditions passaged for three days in liquid media, then tested for production (n=36) with 100 mg/L DTB at their respective IPTG condition. We observed stable production across the replicates at the three lower conditions, with 1-2 outliers at the top (for 0.01 mM IPTG) or bottom (for 0.025 and 0.05 mM IPTG) range of the population. At the high induction conditions (0.1 mM), the tuning phenotype is visible for 8 colonies, corresponding to 22% of the population. From this experiment, we conclude that although the rate of emerging tuning phenotype is reduced in liquid media[19], induction of 0.1 mM is a suitable set-up to explore this phenomenon further.

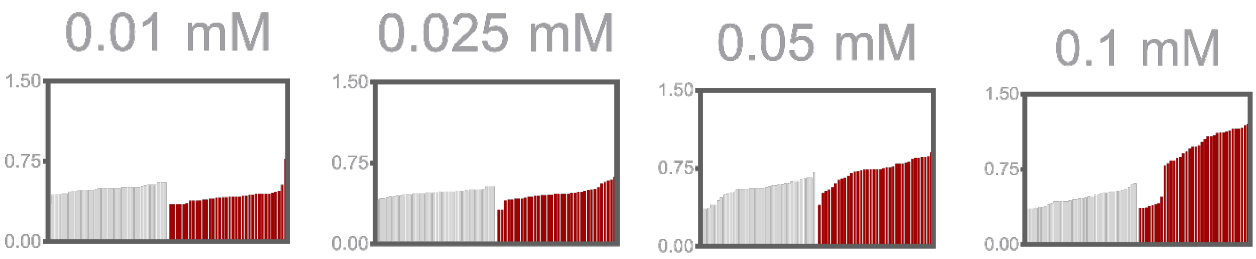


Figure 2: High induction conditions generate the tuning phenotype

A set of BS1574 strains with pBS679 was exposed to the stated IPTG concentrations by passaging for 3 days, then tested for production at 100 mg/L DTB at uninduced (grey) and respective induced (red) conditions. Titers were determined by bioassay. As the plate-to-plate variations of the bioassay increase at the lower limit of detection, all values were normalised to the respective lowest titer data point.

[18] Strain BW25113 with Δbio, expression mutant WT to simulate initial selection strain.
[19] Compared to the 100% false positive rate from the previous selections.

A construct that transcriptionally fuses BioB with a reporter gene

In order to a) study the tuning phenotype more closely and b) generate a counter-screen to re-enable toxicity selections in the future, we build a synthetic construct to transcriptionally fuse the *bioB* sequence with a reporter gene, in this case *sfGFP*.

In the native biotin operon, *bioBFCD* are regulated by the same P_B promoter. Gene *bioB* overlaps with *bioF* by four bases. Replacing the latter with *gfp* generates a transcriptional fusion *bioB-gfp*. Plasmid pBS1058 was constructed in this manner (Figure 3) yet otherwise mirrors the plasmid architecture of pBS679: IPTG-inducible GOI, low copy number origin, additional HindIII site, and ampicillin resistance. To utilise GFP output as a suitable indicator of BioB transcription, we needed to show that plasmid pBS1058 behaves in a parallel manner to pBS679 in small-scale, plates, and shake-flasks.

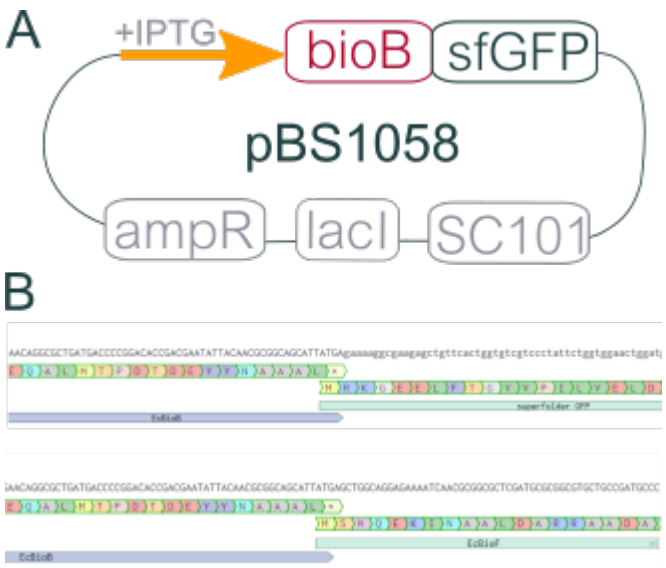


Figure 3: Fusion of BioB and GFP in plasmid pBS1058

A - Plasmid map of pBS1058 showing IPTG-inducible bioB transcriptionally fused to a GFP reporter gene in a plasmid backbone with low copy number (SC101 origin of replication) and conferring ampicillin resistance. HindIII after GOI site not shown.

B - Transcriptional fusion of the *bioB* and *gfp* genes mirrors the native E. coli bio operon for *bioB* and *bioF* sequences. Overlap visualised in Benchling.

Characterisation in small-scale and shake-flask cultivations

In strain BS2163 (BS1574[20] with pBS1058), we characterized behaviour of the fusion plasmid for growth, production, and fluorescence based on IPTG induction (Figure 4) to compare to the activity of pBS679 (Chapter 3). We observed expression toxicity at high induction, no reduction in biotin titer compared to pBS679, and fluorescence mirroring BioB expression. However, the toxicity at high induction differed from pBS679 as it manifested as an increase in lag-time at concentrations of 0.1 mM IPTG and above instead of a reduction in end-OD and cell death[21]. We observed that fluorescence was maximally inducible even when lag-time toxicity was observed, indicating that GFP expression does not confer toxicity. This is confirmed with plasmid pBS683, which only expressed GFP under the T5lac-promoter and shows no toxicity (Figure 6).

[20] Experiments were carried out in the expression mutant WT strain (not FeS mutant strain BS1575) to mimic selection conditions.
[21] The effect of IPTG concentrations above 10 mM were not tested.

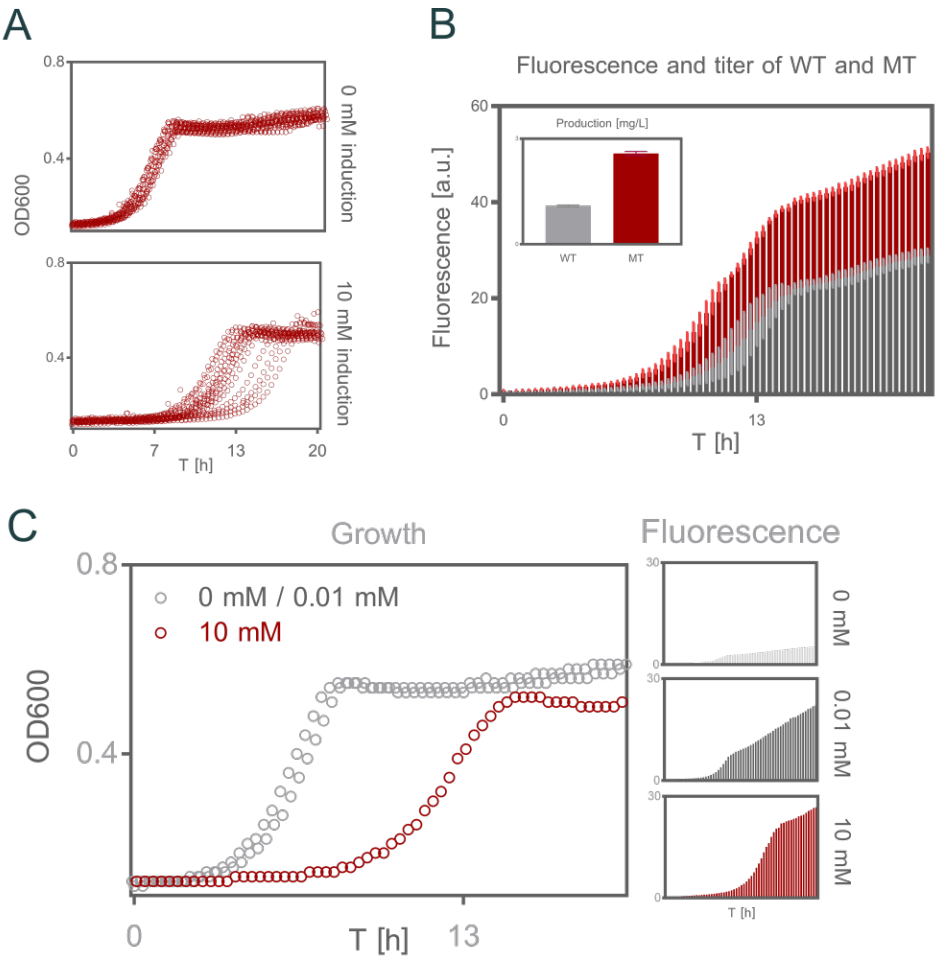


Figure 4: The GFP output of pBS1058 mirrors its BioB transcription

A - Growth curves of BS2163 in the presence of 10, 50, or 100 mg/L DTB at no induction (top) or high induction (10 mM, bottom). Induction has a negative effect on growth by extending the lag-time but not end-OD or growth rates. An increase in lag-time was observed for replicates with 10 mg/L DTB.

B - A known expression mutant increases not only production titers (inset, mutant in red) but also fluorescence at all time points.

C - Comparing growth and fluorescence of BS2163 at three induction conditions indicates a trade-off between these two parameters. The left graph shows growth analogously to A. Again, lag-phase is extended at high induction. A medium induction condition (0.01 mM IPTG) does not lead to a visible phenotypic difference, indicating that growth is not affected by the medium induction conditions. However, the fluorescence output of the strain under no, medium, and high inductions (right boxes) shows that despite the growth deficiency, 10 mM IPTG induces the highest amount of fluorescence following exponential phase whereas the medium induction accumulates fluorescence more slowly and to a lesser degree.

The plasmid was also tested in the expression mutant identified in previous rounds of toxicity selection (Figure 4B). As with pBS679, toxicity of pBS1058 is decreased in this strain. We observe a reduction in lag-time and increase in both fluorescence and titer in the MT strains. We also looked at the effect of DTB on plasmid pBS1058 and observed an *increase* in lag-time at high DTB and IPTG conditions (Figure 4A, bottom). Potentially, availability of the precursor increases activity of the BioB enzyme and therefore increases toxicity. We did not pursue this avenue further. We also characterised plasmid behaviour in shake-flask (Figure 5) to confirmed scalability. From the data we conclude that plasmid pBS1058 links transcription of *bioB* and *gfp.*

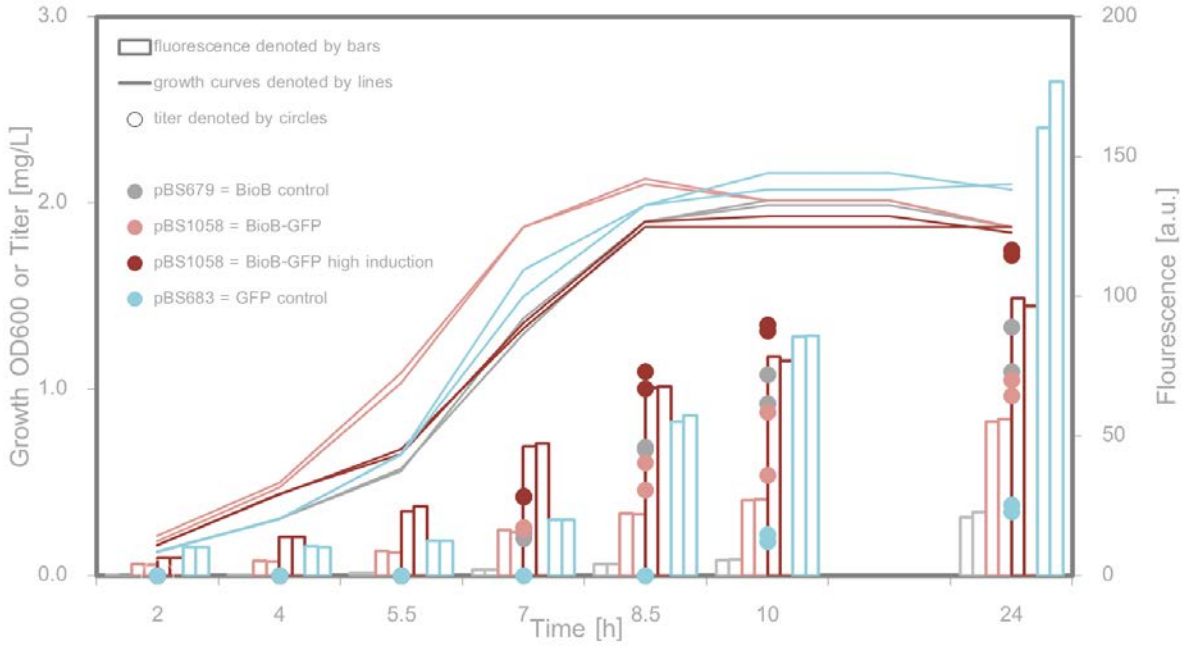


Figure 5: Plasmid behaviour scales to shakeflasks

A shake-flask experiment (n=2) was carried out to compare pBS697 (grey), pBS1058 at low (pink) and high (red) induction, and control pBS683 (blue). Shown is OD (lines), GFP fluorescence (bars), and biotin production (dots) as determined with the bioassay. Note that due to high fluorescence, the GFP values of pBS683 are 10% of their total values.

After mutagenic conditions, the tool sieves out false positives

To generate diversity, strain BS2163 (from here on referred to as BioB-GFP plasmid) was passaged for three days in minimal MOPS containing ampicillin, DTB (1 µM), and high levels of IPTG (10 mM). As controls, strains BS2385 (from

here on referred to as BioB plasmid) and BS2186 (from here on referred to as GFP plasmid) were also passaged[22].

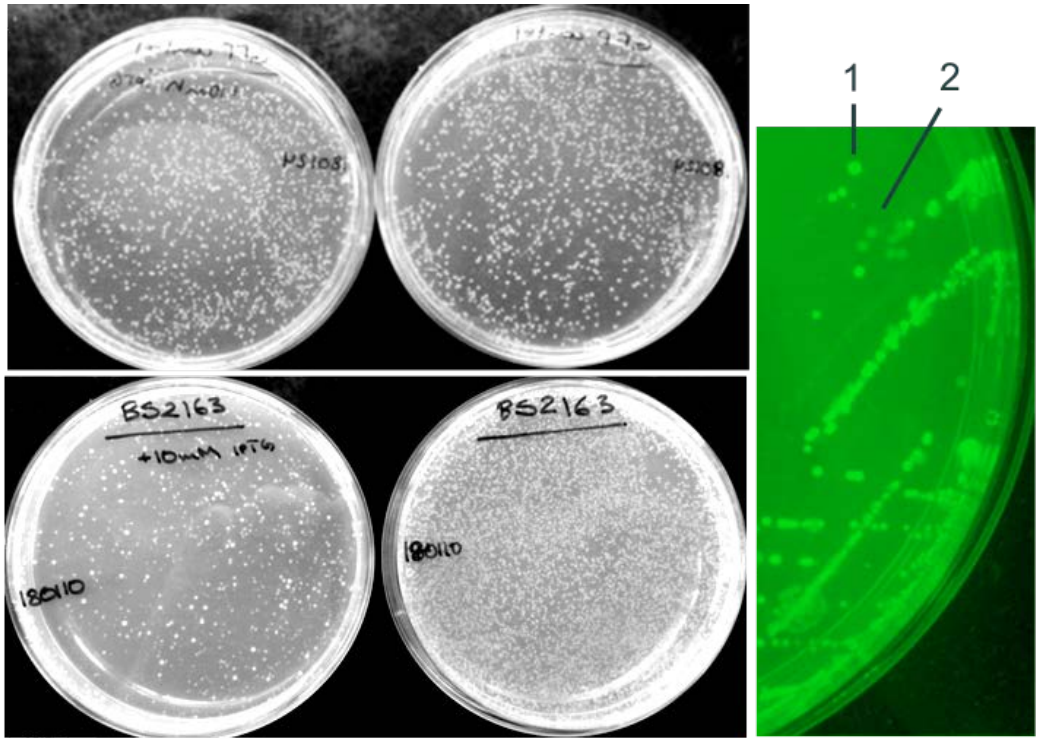


Figure 6: Colony formation after toxicity of high induction

Colony formation after IPTG induction of strain harbouring the GFP plasmid (top row) or BioB-GFP (bottom row). Both strains were exposed to either 0 or 10 mM IPTG for 3 days before plating on no induction (right) or 10 mM IPTG (left). Based on colony size, phenotypic diversity is discernible for BS2163 both not BS2186. Diversity can also be seen for individual colony behaviour under blue light (right photo): the majority of colonies show fluorescence visible to the eye (indicated by 1) whereas ~30% of colonies show no brightness (no expression of GFP, indicated by 2).

Strains were plated out on minimal agar plates at high induction (10 mM for BioB-GFP fusion and GFP control, 0.1 mM for BioB plasmid) conditions and incubated at 37 °C for 48 h (Figure 6). Diverse colony morphology was discernible for BioB plasmids and BioB-GFP fusion. The latter also showed where both highly fluorescent and non-fluorescent colonies. The quantify this diversity further, 92 colonies of BS2163 were picked along with 46 colonies of the BioB and GFP controls. Further, it was noted whether the BS2163 colony in each position had lost fluorescence in the plating condition (~24% observed loss of fluorescence, Figure 6). A production test was carried out for the three strains to determine gain

[22] In each case, the BioB plasmid was induced at 0.1 mM IPTG. The GFP plasmid was supplemented with 1 µM BTN.

of phenotypic diversity based on three parameters under two conditions: growth, production, and fluorescence at leaky and high expression (Figure 7). Comparison of these values indicate that, taking end-OD as a determinant of fitness, all tested strains were amenable to the production conditions. Secondly, fluorescence was highly inducible for the GFP plasmid and as expression of GFP in this plasmid construct is not toxic, there is little metabolic pressure to affect expression levels or result in tuning. No fluorescence above background levels was observed for the BioB strain.

We separated the BioB-GFP plasmid strains data into two sets, with or without observable GFP phenotype on the plate (Figure 8). Although the majority of non-observable strains showed some fluorescence on the plate-reader, the maximum emission values reached no more than 25% of the expected values. This was mirrored by biotin titers: numbers were less than 25% of expected values. There was no observable correlation between GFP and titer of the non-fluorescent colonies. We compared titer and fluorescence of the observable fluorescence colony set and observed a robust correlation ($n=70$, $R^2=0.7$) even after exposure to selective conditions.

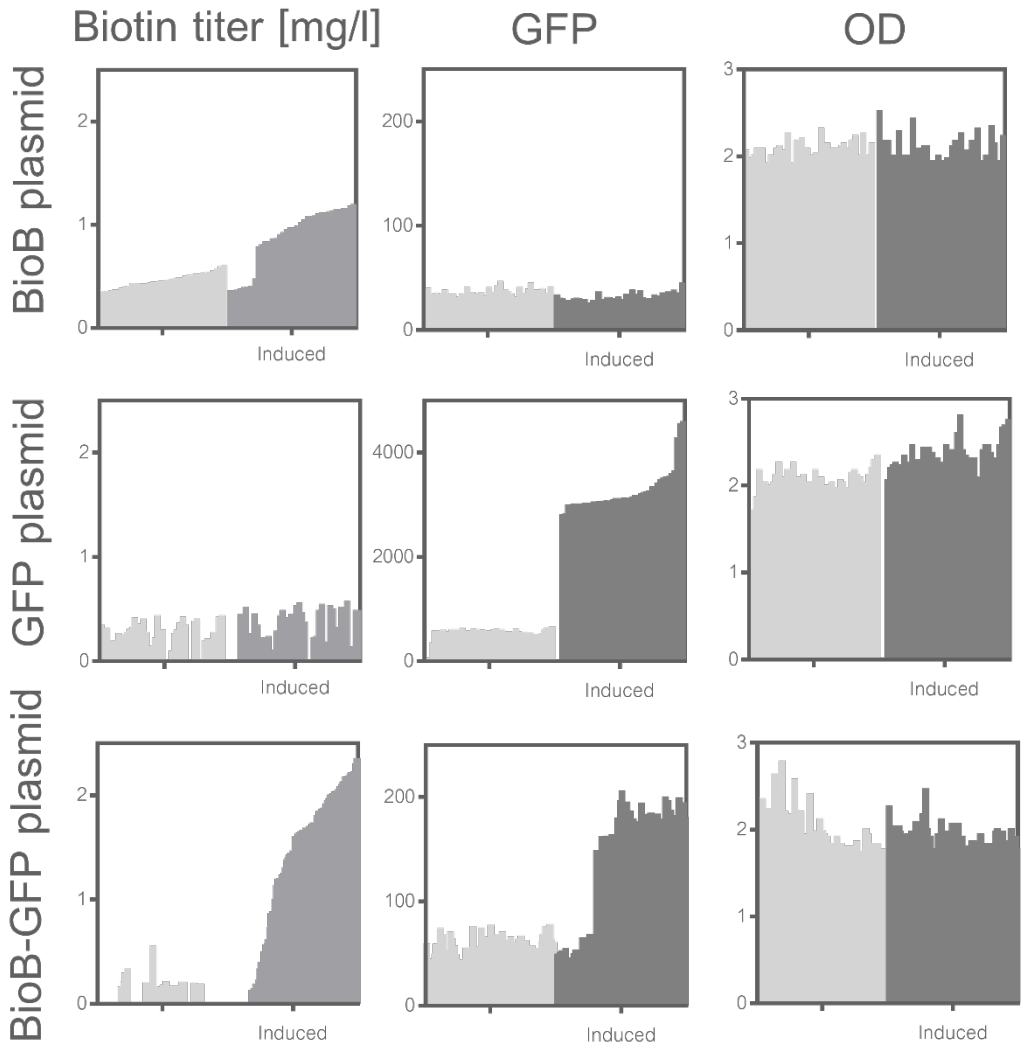


Figure 7: Diversity in plasmid behaviour upon induction

OD, fluorescence, and titer were calculated for each strain under induced (right) and un-induced conditions (left). To aid the eye, BioB and BioB-GFP plasmid data were sorted lowest to highest based on induced titer, GFP plasmid was sorted based on induced fluorescence. Note that i) GFP plasmid contained 1 µM biotin in the production media, which is detected on some of the samples. ii) BioB plasmid is tested for production at 0.1 mM IPTG, which is not optimal for this plasmid.

In order to quantify the efficiency of the constructed counter-screen, we simulated a two scenarios of toxicity selection on the generated data (Figure 7 & 8). In each case, a random colony of BS2163 is picked and tested for production after exposure to toxicity conditions, i.e. high IPTG induction. Given the small number of colonies characterised in this data set, we assume no selection hits (variants with a higher production than the WT) and therefore divide the producers into

categories of poor or expected titers. A poor producer is here defined as producing less than 30% of the expected titer.

Scenario 1 (Figure 8, top boxes) does not take the GFP behaviour of BS2163 into account. Based on the data shown (n=92), 29 colonies produced poorly, i.e. less than 0.7 mg/L at 0.1 mM induction. The likelihood when picking a random colony to obtain a poor producer is therefore 32%. In the second Scenario (Figure 8, bottom boxes) the GFP behaviour of BS2163 is taken into account. Therefore the colonies are pre-screened for fluorescence on selection plates prior to testing for production. This rules out 22 colonies of the total screen, leaving 70 colonies. Only 7 of these were poor producers such that the likelihood of randomly picking a poor producer is 10%. The group of library members characterised by high fluorescence were significantly different than the group as a whole (p=0.0134, non-paired Mann Whitney test, calculated in prism). From these data, we conclude that the BioB-GFP fusion is a successful tool to counter-screen for the tuning phenotype: the total number of colonies requiring production characterisation is reduced by over one third, and the likelihood of a tuned phenotype when utilising the screen is reduced 3-fold.

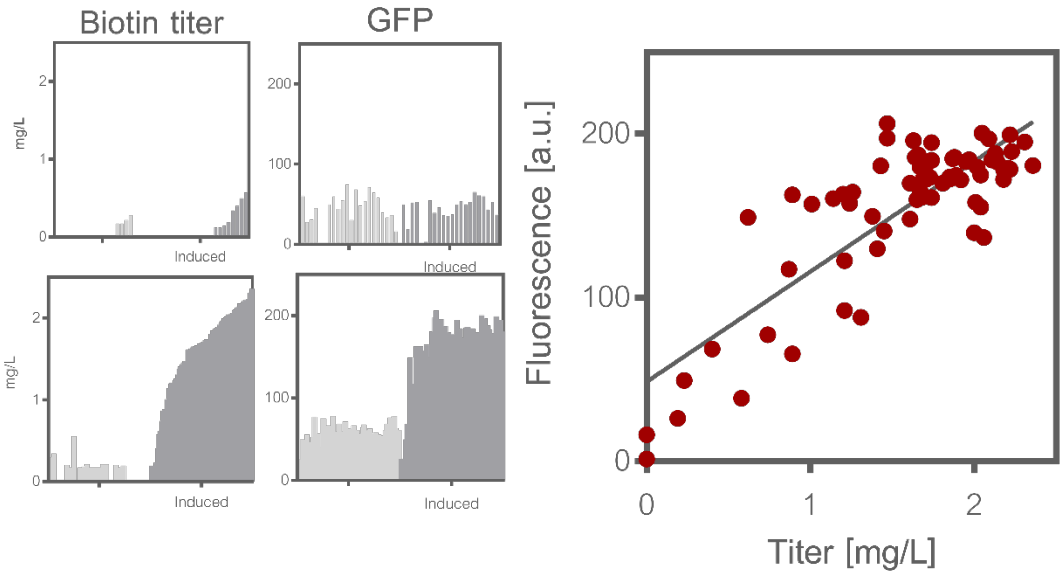


Figure 8: The fusion plasmid is a robust tool for selection

Left - Fluorescence and titer for BS2163 strain under induced and un-induced conditions based on whether colony fluorescence was visible (bottom) or not (top). The top boxes are considered “tuned”, i.e. IPTG inert. The colonies visualized in the bottom boxes

produce and fluoresce as expected. To aid the eye, plasmid data were sorted lowest to highest based on induced titer.

Right – Correlation of titer and fluorescence of non-tuned strains indicate high robustness of the fusion (R^2 = 0.68, n=70).

DISCUSSION

Loss of functionality arises in most biological systems when a “production load” is present, due to factors including protein misfolding, the cost of DNA and enzyme synthesis, toxicity of products, or misbalanced use of cofactors (3). The latter is pertinent to our system due to observed toxicity of BioB expression, which we hypothesise could be due to i) drain on endogenous iron-sulfur pools or ii) ROS generation during enzyme regeneration process. In our set-up, enzyme expression is coupled to an extracellular inducer, leading to loss of IPTG sensitivity as a first-response mutation.

In previous work on loss-of-function, alleviation of production burden is shown as a highly effective countermeasure. The growth rate cost of plasmids could be diminished by expressing gene on the plasmid of choice to direct carbon flux into additional building block biosynthesis as a counter-balance for production burden (8). A second countermeasure is to reduce the mutation rate of the strain by deleting less reliable polymerases (9). Yet these measures are not feasible in our system, as we were relying on the toxicity of production and *E. coli*’s natural mutation rate to identify potential overproducers. To construct a compromise counter-screen for loss-of-function, we engineered an internal BioB monitor by transcriptionally but not translationally fusing the gene to *sfGFP*. This complemented our previous analytical methods of extracellular biotin measurement via bioassay and protein determination by proteomics by shedding light on BioB transcription in a semi-quantitative manner. Further, the fusion approach allowed us to study BioB toxicity under various conditions as GFP expression generated no additional burden on the system at high induction. This translates into minimal pressure to evolve the monitor plasmid, compared to antibiotic resistance of kill-switch systems with a high rate of escape.

Upon characterisation, we observed that the BioB-GFP fusion plasmid was *more* stable to IPTG induction and had a high production capacity which surpassed that of pBS679 in some conditions. It is not clear why the addition of GFP

decreases toxicity of BioB expression. Hypotheses include stabilisation of BioB mRNA, gradiation of expression at lower concentrations, or an increased anaerobic environment due to the oxygen requirement of GFP maturation. Unfortunately, the stabilization by GFP has a result that the original toxicity screen using BioB induction could not be carried out with the fusion plasmid, as it was not possible to induce plasmid pBS1058 to toxicity. Nevertheless the plasmid is of high utility in other selections around the BioB gene, such as co-factor engineering via MAGE. Alternatively, the construct could be generated on a higher copy-number plasmid to reinstate BioB toxicity at high induction. Finally, the correlation between GFP and biotin titers was sufficiently robust that strain fluorescence could act not only as a counter-screen for tuning, but act as a screening tool itself, despite some changes in expression kinetics. Overall, we are confident that the BioB monitor plasmid described in this chapter is a suitable addition to the selection box for identifying biotin overproducers.

BIBLIOGRAPHY

[1] Tominaga M, Kawai-Noma S, Kawagishi I, Sowa Y, Saito K, Umeno D. **Liquid-based iterative recombineering method tolerant to counter-selection escapes.** *PLoS ONE* 2015; **10(3):**1–18.

[2] Sieben M, Steinhorn G, Müller C, Fuchs S, Ann Chin L, Regestein L, Büchs J. **Testing plasmid stability of Escherichia coli using the Continuously Operated Shaken BIOreactor System**. *Biotechnology Progress* 2016; **32(6):**1418–25.

[3] Rugbjerg P, Myling-Petersen N, Porse A, Sommer M. **Diverse genetic error modes constrain large-scale bio-based production.** *Nature Communications* 2018; **9(1):787**.

[4] Casanova M, Pasotti L, Zucca S, Politi N, Massaiu I, Calvio C, et al. **A BioBrickTM-compatible vector for allelic replacement using the xylE gene as selection marker.** *Biological Procedures* 2016; **18(1):**37.

[5] Carrico ZM, Strobel KL, Atreya ME, Clark DS, Francis MB. **Simultaneous selection and counter-selection for the directed evolution of proteases in *E. coli* using a cytoplasmic anchoring strategy.** *Biotechnology and Bioengineering* 2016; **113(6):**1187–93.

[6] Aertsen A, Houdt R Van, Michiels CW. **Construction and use of an stx1 transcriptional fusion to gfp.** *FEMS Microbiology Letters* 2005; **245(1):**73–7.

[7] Rufián JS, López-Márquez D, López-Pagán N, Grant M, Ruiz-Albert J, Beuzón CR. **Generating chromosome-located transcriptional fusions to fluorescent proteins for single-cell gene expression analysis in *Pseudomonas syringae.*** In *Host-Pathogen Interactions: Methods and Protocols*. Springer, New York; 2018. p. 183–99.

[8] Flores S, De Anda-Herrera R, Gosset G, Bolívar FG**. Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway.** *Biotechnology and Bioengineering* 2004; **87(4):**485–94.

[9] Csörgo B, Fehér T, Tímár E, Blattner FR, Pósfai G. **Low-mutation-rate, reduced-genome *Escherichia coli:* An improved host for faithful maintenance of engineered genetic constructs**. *Microbial Cell Factories* 2012; **11:**1–13

ANNEX

Media and stock solutions

10 MOPS

Up to a total volume of 300 mL, 83.72 g of MOPS buffer and 7.17 g tricine were dissolved in MQ and buffered to pH = 7.4 using 10 M KOH (90% Reagent Grade, Sigma Aldrich) at a final volume of 440 mL. Fresh iron sulfate solution is made by dissolved 0.028 g $FeSO_4 \bullet 7H_2O$ in 10 mL MQ and added. To a final volume of 1l, salts are added: 50 m 1.9 M NH_4Cl, 10 mL 0.276 M K_2SO_4, 5 µL 1M $CaCl_2 \bullet 2H_2O$, 100 mL 5M NaCl. Of the micronutrient stock, 0.2 mL is added. The media stock is sterile-filtered.

10x M9 salts

In a total volume of 1 L MQ, 68 g Na_2HPO_4 * $7H_2O$, 30 g KH_2PO_4 , 5 g NaCl, and 10 g NH_4Cl are dissolved and sterile-filtered.

Vitamin solution

Four solutions of 25 mL each of 0.238 g calcium pantothenate (Sigma-Aldrich) in MQ, 0.068 *p*-aminobenzoic acid (Sigma-Aldrich), 0.069 *p*-hydroxybenzoic acid (Sigma-Aldrich), and 0.077 2,3-dihydroxybenzoic acid (Sigma-Aldrich) all in KOH, were mixed, sterile-filtered, and aliquoted.

Trace element solution

In a total of 20 mL MQ, 11 mg $ZnSO_4 \cdot 7H_2O$, 10 mg $CuSO_4 \cdot 5H_2O$, 6 mg $MnSO_4 \cdot H_2O$, 0.835 g $FeCl_3 \cdot 6H_2O$, 11 mg EDTA disodium, 11 mg $CoCl_2 \cdot 6H_2O$, and 40 mg $CaCl_2 \cdot 2H_2O$ are dissolved. Mixture can be heated to 60 °C if needed to aid solubilization. Solution is sterile-filtered

10x TAE buffer

In a total volume of 1L, 242 g Tris base (TRIZMA from Sigma-Aldrich), 57.1 mL acetic acid, and 100 mL 0.5M EDTA (Sigma) are mixed.

Strains mentioned in this thesis

Strain name	Plasmid	Genotype
BS1011	-	BW25113 Δ*bioB*
BS1013 BW25113	-	*lacI*$^{+}$*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR*514 Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$ *rph-1* Δ*(araB–D)567* Δ*(rhaD–B)568* Δ*lacZ4787*(::*rrnB-3*) *hsdR514 rph-1*
BS1020		BW25113 ΔyigM
BS1093	pBS451	BW25113 Δ*bioB*
BS1574	-	BW25113 Δ*bio*
BS1575	-	BW25113 Δ*bio* Fe-S mutation
BS1601		BW25113 Δ*bio* BirA I207S
BS1613	pBS910 (CU.4)	BW25113 Δ*bio,* Fe-S mutation
BS1874	pBS679	BW25113 Δ*bio* BirA I207S
BS2239	pBS679	BW25113 Δ*bio* Fe-S mutation, ?
BS2240	pBS679	BW25113 Δ*bio* Fe-S mutation, ?
BS2151	pBS679	BW25113 Δ*bio* Fe-S mutation, ?
BS2163	pBS1058	BW25113 Δ*bio*
BS2186	pBS683	BW25113 Δ*bio*
BS2385	pBS679	BW25113 Δ*bio*
sAS278		29ΔglkΔpts1ΔmanZ1ΔyigM
Top10		F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^{R}) endA1 λ^{-}

Oligos

oBS020	CTGTACAAATGATGATCTAGAGGCATCAAATAAAACGAAAGGCTC
oBS021	TTCGCCTTTACGCATAAGCTTAGAACCGCCTCCAG
oBS768	ACCCCGGGGCUTCTCCAAAACGTGTTTTTTGTT
oBS789	ACTACCAUGCCACAAACCAGCTAAGGAGGTAAATATGAGCTGGCAGGAGAAAATCAAC
oBS886	TGATGTTGGCAGTTTTCTCG
oBS889	AATGGTTTCTTAGACGTCGGAATTGCCAGCGATTTCCCTCGAGGTGAAGACGAAAG
oBS1007	GTCCCACCGCTTACCTCCTATTACGTCACTAATAGCTAAGGAGGTAAATATGAATAACATCTGGTGGCAGACC
oBS1105	ATGCGTGAGATCCAGCTT
oBS1106	AAGATGAUCTTCTTGAGATCGTTTTGGTCT
oBS1107	ACCCCGUAGGCGGCAAAGCCGTTTTTC
oBS1159	ACTGAGCUAGCTGTAAACTAGCATAACCCCTTGGG
oBS1160	ATGTGTTTGUTTATTTTTCTAAATACATTCAAAT
oBS1161	ACAAACACAUTTCCCCGAAAAGTGCCACCTG
oBS1162	ATCTCAAACUCAGGAGAGCGTTCACCGACAA
oBS0768	AGCGGTTGAUGCGTAAAGGCGAAGAGCTGTT
oBS0769	ATTGGGGUGCCTAATGAGTGAGCTAA
oBS1105	ATCATCTUAAACGCCAGCAACGCGGCC
oBS1172	AGGGGCGGGGUTTTTTTTTAGTGCTTCATGTGGCAGGAGAA
oBS1173	ATTCCTCUACGCCGGACGCATCG
oBS1159	ACTGAGCUAGCTGTAAACTAGCATAACCCCTTGGG
oBS1160	ATGTGTTTGUTTATTTTTCTAAATACATTCAAAT
oBS1279	ACCAAGCACUAGGGACAGTAAGACGGGTA
oBS1280	ATCCCACUCCCCCTAGAGGCATCAAA
oBS1283	AGCCATATUTACCTCCTTAGCTCTAGAG
oBS1284	AGTGCTTGGUGATAGGGACTCGAGGTGAA
oBS1435	AACCAATUGAGGTCCGAGGGAATGAATGAATCTATTATAGGTACAAAAAGATGCG
oBS1440	AGTCTACCUTCAAGCGCCCGTTCACCATGCGTAAAGCAATCAG
oBS1587	ACAAGCTUTCCCACTCCCCCTAGAGGC
oBS1590	AAGCTTGUTCATCATTTGTACAGTTCATCCATACCATGCG
oBS1591	AGAGATACUAGTTAGAGAACCTCCGATTTGAATC
oBS2236	ACTGCCAUCCACAGAACCTAAGAGCCGGATGATTAATTGTCAACACCATTTGCCAGCTGG
oBS2237	ACTGCCAUCCACACACGCTAAGAGCCGGATGATTAATTGTCAACACCATTTGCCAGCTGG
oBS2374	AGCATTAUGAGAAAAGGCGAAGAGCT
oBS2375	AGCTTGTTUTGTACAGTTCATCCATACCATG
moBS314	GCTTCCTGCAGCGTGATCCACCCCTGATTAACGACACTCTCTTCAACACGGCGCATTGCCATGTTGCTCCCGGCTCCAATGACTATTTGC

Contributions

Introduction

Texts were written by JB. Luisa Gronenberg and Hans Jasper Genee provided comments. David Lennox assisted with translating Danish abstract.

Chapter 1

Texts written by JB, LG, HJG, and Morten Sommer edited and provided input. Images were made under initial supervision of HJG.

Chapter 2

Texts written by JB. LG, HJG and Lukas Herwig provided comments. Biosensor construct V1 was constructed by Lasse Holm Lauridsen. Nanoliter protocols were performed under supervision of LH. LH also generated data used in Figure 8B. Proof-of-concept library was constructed by Nils Myling-Petersen. JB characterized and sequenced the constructs. Remaining work performed by JB, supervised by LG and HJG.

Chapter 3

Texts written by JB. LG and HJG provided comments. Experimental concepts developed under supervision of LG and HJG. Plasmids pBS511 and pBS1406 constructed by Bo Salomonson, pBS479 by Anne Pihl Bali. Expression mutant identified by APB. Metagenomic libraries were generated in the Sommer lab and kindly provided. Precursor library generated by NMP. Remaining work performed by JB, supervised by LG and HJG.

Chapter 4

Texts written by JB. LG and HJG provided comments. Experimental concepts developed under supervision of LG and HJG. Plasmid pBS683 constructed by Bo Salomonson. Remaining work performed by JB, supervised by LG and HJG.